

(11)Publication number : 2003-333599  
(43)Date of publication of application : 21.11.2003

H04N 7/32

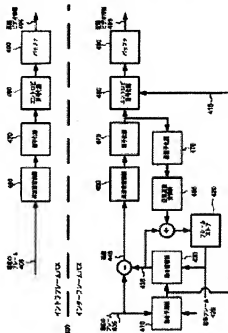
(71)Applicant : MICROSOFT CORP

(72)Inventor : SRINIVASAN SRIDHAR

Priority country : US  
US

(57)Abstract:

**SOLUTION:** For example, the video encoder or the video decoder uses an intermediate pixel values with a dynamic range (in the unit of bits) wider than a final range to calculate a pixel value at a sub-pixel sample position (e.g. an intermediate value of 16 bits and an output value in 8 bits). The encoder or the decoder at least partially defers shifting from a first stage to a second stage or can skip clamping in the first stage or in other intermediate stages for multi-stage interpolation.



(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2003-333599

(P2003-333599A)

(43) 公開日 平成15年11月21日 (2003. 11. 21)

(51) Int.Cl.<sup>7</sup>

H 0 4 N 7/32

識別記号

F I

H 0 4 N 7/137

データベース (参考)

Z 5 C 0 5 9

審査請求 有 請求項の数41 O L (全 29 頁)

(21) 出願番号 特開2003-107081 (P2003-107081)

(22) 出願日 平成15年4月10日 (2003. 4. 10)

(31) 優先権主張番号 6 0 / 3 7 1, 8 6 0

(32) 優先日 平成14年4月10日 (2002. 4. 10)

(33) 優先権主張国 米国 (US)

(31) 優先権主張番号 1 0 / 3 8 2 3 1 1

(32) 優先日 平成15年3月4日 (2003. 3. 4)

(33) 優先権主張国 米国 (US)

(71) 出願人 391055933

マイクロソフト コーポレーション  
MICROSOFT CORPORATI  
ONアメリカ合衆国 ワシントン州 98052-  
6399 レッドモンド ワン マイクロソ  
フト ウェイ (番地なし)

(74) 代理人 100077481

弁理士 谷 義一 (外1名)

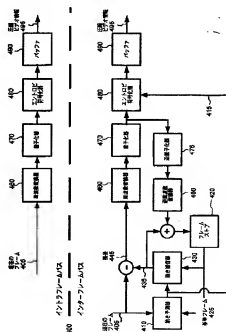
最終頁に続く

(54) 【発明の名称】 動き予測および動き補償におけるサブピクセル補間

## (57) 【要約】

【課題】 動き予測および動き補償のさまざまな技法およびツール (たとえば、ビデオエンコーダ/デコーダ内の) を提供すること。

【解決手段】 例えば、ビデオエンコーダまたはビデオデコーダは、最終的な値より広いダイナミックレンジ (ビット単位) を有する中間ピクセル値を使用して、サブピクセルサンプル位置のピクセル値を計算する (例えば、16ビットの中間値および8ビットの出力値)。エンコーダまたはデコーダは、少なくとも部分的に、第1ステージから第2ステージへのシフトを延期するか、または複数ステージ補間の第1ステージまたは他の中間ステージにおけるクランプをスキップすることができる。



1

【特許請求の範囲】

【請求項 1】 コンピュータシステムにおいて、コンピュータで実施される方法であって、  
1または複数の近似サブサンプル位置における1または複数の中間値の少なくとも一部に基づいて、特定のサブサンプル位置の最終的な値を計算することを備え、  
前記最終的な値は、 $x$  ビットの最終的なダイナミックレンジを有し、前記1または複数の中間値は、 $y$  ビットの間ダイナミックレンジを有し、 $y$  が  $x$  より大きいことを特徴とする方法。

【請求項 2】  $y$  は、16 ビット以下であることを特徴とする請求項 1 に記載の方法。

【請求項 3】  $x$  は、8 ビットであることを特徴とする請求項 1 に記載の方法。

【請求項 4】 前記計算することは、複数ステージ補間のより前のステージから複数ステージ補間のより後のステージへのビットシフトを延期することを含み、前記延期されるビットシフトは、サブサンプルシフトの異なる度合の量に応じて変ることを特徴とする請求項 1 に記載の方法。

【請求項 5】 前記計算することは、複数ステージ補間のより前のステージから複数ステージ補間のより後のステージへのビットシフトを延期することを含み、前記延期されるビットシフトは、前記複数ステージ補間の前記より前のステージおよび前記より後のステージに関する拡張係数に基づいて変ることを特徴とする請求項 1 に記載の方法。

【請求項 6】 前記特定のサブサンプル位置および前記近似サブサンプル位置は、基準ビデオフレーム内であり、前記最終的な値は、最終的なピクセル値であり、前記中間値は、中間ピクセル値であることを特徴とする請求項 1 に記載の方法。

【請求項 7】 前記特定のサブサンプル位置は、前記基準ビデオフレームの少なくとも1つの次元において1/2ピクセルシフトに配置されることを特徴とする請求項 6 に記載の方法。

【請求項 8】 前記特定のサブサンプル位置は、前記基準ビデオフレームの少なくとも1つの次元において1/4ピクセルシフトまたは3/4ピクセルシフトに配置されることを特徴とする請求項 6 に記載の方法。

【請求項 9】 前記特定のサブサンプル位置は、前記基準ビデオフレームの2つの次元においてサブピクセルシフトに配置されることを特徴とする請求項 6 に記載の方法。

【請求項 10】 前記計算の第1ステージは、前記1または複数の中間値をもたらすことを特徴とする請求項 1 に記載の方法。

【請求項 11】 前記第1ステージは、垂直フィルタリングを備え、前記第2ステージは、水平フィルタリング

2

を備えたことを特徴とする請求項 10 に記載の方法。

【請求項 12】 1または複数の追加のサブサンプル位置の各々について前記計算を繰り返すことをさらに含むことを特徴とする請求項 1 に記載の方法。

【請求項 13】 コンピュータシステムに、ビデオ符号化中に請求項 1 に記載の方法を実行させるコンピュータ実行可能命令を記憶することを特徴とするコンピュータ読み取り可能媒体。

【請求項 14】 コンピュータシステムに、ビデオ復号化中に請求項 1 に記載の方法を実行させるコンピュータ実行可能命令を記憶することを特徴とするコンピュータ読み取り可能媒体。

【請求項 15】 コンピュータシステムにおいて、複数のビデオフレーム間の時間的冗長性を利用してコンピュータで実施される方法であって、

複数ステージの補間を使用して、前記複数のビデオフレームの基準ビデオフレームにおける小数ピクセル位置の最終的なピクセル値を計算することを備え、

前記複数のステージは、少なくとも第1ステージおよび最終ステージを含み、前記計算することは、前記第1ステージ中の1または複数の中間ピクセルのクランプをスキップすることを含み、前記計算することは、前記最終ステージ中の前記基準ビデオフレームにおける前記小数ピクセル位置の前記最終的なピクセル値をクランプすることを含むことを特徴とする方法。

【請求項 16】 前記小数ピクセル位置は、前記基準ビデオフレームの第1の次元における1/4ピクセルシフト、1/2ピクセルシフト、または3/4ピクセルシフトであり、前記基準ビデオフレームの第2の次元における1/4ピクセルシフト、1/2ピクセルシフト、または3/4ピクセルシフトであることを特徴とする請求項 15 に記載の方法。

【請求項 17】 前記複数のステージの1または複数のステージは、近似双三次フィルタを使用することを特徴とする請求項 15 に記載の方法。

【請求項 18】 前記計算することは、前記ピクセル値の精度を高めるために前記第1ステージから前記最終ステージへのビットシフトを延期することをさらに含むことを特徴とする請求項 15 に記載の方法。

【請求項 19】 前記延期されるビットシフトは、小数ピクセルシフトの異なる度合の量に応じて変ることを特徴とする請求項 18 に記載の方法。

【請求項 20】 前記延期されるビットシフトは、前記第1ステージおよび前記最終ステージに導入される拡張の量に応じて変ることを特徴とする請求項 18 に記載の方法。

【請求項 21】 前記第1ステージは、第1方向での第1の丸めを含み、後続ステージは、前記第1位置と反対の第2方向での第2の丸めを含み、丸め制御パラメータは、前記第1方向および前記第2方向を制御することを

50

特徴とする請求項 15 に記載の方法。

【請求項 22】 コンピュータシステムに、ビデオ符号化中に請求項 15 に記載の方法を実行させるコンピュータ実行可能命令を記憶することを特徴とするコンピュータ読み取り可能媒体。

【請求項 23】 コンピュータシステムに、ビデオ復号化中に請求項 15 に記載の方法を実行させるコンピュータ実行可能命令を記憶することを特徴とするコンピュータ読み取り可能媒体。

【請求項 24】 コンピュータシステムにおいて、コンピュータで実施される方法であって、特定の小数サンプル位置で、複数ステージの補間を使用して値を計算することを備え、

該計算することは、前記複数ステージの第 1 ステージから前記複数ステージの第 2 ステージへのビットシフトを延期することを含み、前記値の精度を高めることを特徴とする方法。

【請求項 25】 前記計算される値は、最終的な値であり、前記計算することは、

前記第 1 ステージにおいて、近似小数サンプル位置の複数の中間値を計算し、1 または複数のビットの第 1 シフト量だけ前記複数の中間値の各々をビットシフトすることと前記第 2 ステージにおいて、前記複数の中間値の少なくとも一部に基づいて前記最終的な値を計算し、前記第 2 ステージの正規化係数の 2 を底とする対数より大きい第 2 シフト量だけ前記最終的な値をビットシフトすることを備えたことを特徴とする請求項 24 に記載の方法。

【請求項 26】 前記第 2 シフト量は、前記第 1 シフト量より大きいことを特徴とする請求項 25 に記載の方法。

【請求項 27】 前記第 2 シフト量は、7 ビットであることを特徴とする請求項 25 に記載の方法。

【請求項 28】 前記複数の中間値は、前記第 1 シフト量による前記ビットシフトの前および前記第 1 シフト量による前記ビットシフトの後の両方で、入力値のダイナミックレンジより広い中間値のダイナミックレンジを有することを特徴とする請求項 25 に記載の方法。

【請求項 29】 前記第 1 シフト量による前記ビットシフトは、前記第 2 ステージにおいてワード境界内にとどまるダイナミックレンジを保つことを特徴とする請求項 25 に記載の方法。

【請求項 30】 前記第 2 シフト量による前記ビットシフトの後には、前記最終的な値は、 $x$  個の仮数ビットおよび 1 個の符号ビットを含むことを特徴とする請求項 25 に記載の方法。

【請求項 31】 前記延期されるビットシフトは、小数サンプルシフトの異なる度合の量に応じて変えることを特徴とする請求項 24 に記載の方法。

【請求項 32】 前記延期されるビットシフトは、前記

複数ステージの補間で使用される 1 または複数の補間フィルタの拡大係数に基づいて変えることを特徴とする請求項 24 に記載の方法。

【請求項 33】 前記計算することは、近似双三次フィルタを使用することを特徴とする請求項 24 に記載の方法。

【請求項 34】 前記特定の小数サンプル位置は、基準ビデオフレーム内であり、前記最終的な値は、最終的なピクセル値であることを特徴とする請求項 24 に記載の方法。

【請求項 35】 1 または複数の追加の小数サンプル位置の各々について前記計算を繰り返すことをさらに含むことを特徴とする請求項 24 に記載の方法。

【請求項 36】 コンピュータシステムに、ビデオ符号化中に請求項 24 に記載の方法を実行させるコンピュータ実行可能命令を記憶することを特徴とするコンピュータ読み取り可能媒体。

【請求項 37】 コンピュータシステムに、ビデオ復号化中に請求項 24 に記載の方法を実行させるコンピュータ実行可能命令を記憶することを特徴とするコンピュータ読み取り可能媒体。

【請求項 38】 基準ビデオフレームの中間サブピクセル位置における複数の中間ピクセル値を計算する手段であって、前記複数の中間ピクセル値の各々は、前記基準ビデオフレームの整数ピクセル位置における複数の整数ピクセル値の間で第 1 の次元に沿って補間することにより計算される手段と、

中間サブピクセル位置における前記複数の中間ピクセル値の間で第 2 の次元に沿って補間することによって、前記基準ビデオフレームの最終的なサブピクセル位置の最終的なピクセル値を計算する手段であって、前記最終的なピクセル値は、 $x$  ビットの最終的なダイナミックレンジを有し、前記複数の中間ピクセル値の各々は、 $x$  ビットより広い中間ダイナミックレンジを有する手段とを備えたことを特徴とするシステム。

【請求項 39】 前記最終的なピクセル値を計算する手段は、前記最終的な値をクランプする手段を含み、前記複数の中間ピクセル値を計算する手段は、前記複数の中間ピクセル値をクランプする手段を含まないことを特徴とする請求項 38 に記載のシステム。

【請求項 40】 前記複数の中間ピクセル値を計算する手段は、第 1 シフト量によりビットシフトする手段を含み、前記最終的なピクセル値を計算する手段は、前記最終的なピクセル値を計算するために、正規化係数の 2 を底とする対数より大きい第 2 シフト量だけビットシフトする手段を含むことを特徴とする請求項 38 に記載のシステム。

【請求項 41】 前記第 2 シフト量は、可変量だけ前記第 1 シフト量より大きいことを特徴とする請求項 40 に記載のシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】ビデオコーディング/デコーディングアプリケーションにおける動き予測/補償でのサブピクセルシフトの技法およびワークが開示されている。例えば、ビデオエンコーダおよびビデオデコーダは、基準フレームのサブピクセル位置においてピクセル値について複数ステージの補間を実行する時に、ビットシフトを延滞する。

【0002】

【従来の技術】デジタルビデオは、大量のストレージおよび伝送容量を消費する。通常の生のデジタルビデオシーケンスには、毎秒15フレームまたは30フレームが含まれる。各フレームに、1万個または10万個のピクセル（ペルとも称する）が含まれる可能性がある。各ピクセルは、画像の小さい要素を表す。生の形では、コンピュータは、一般に24ビットを用いて1ピクセルを表す。例えば、1つのピクセルに、ピクセルのグレイスケール成分を定義する8ビット輝度値（明度値とも称する）と、ピクセルの色成分を定義する2つの8ビットクロミナンス値（クロマ値とも称する）が含まれる場合がある。従って、典型的な生のデジタルビデオシーケンスのビット毎秒またはビットレート（秒あたりのビット数）は、毎秒50万ビット以上になる可能性がある。

【0003】多くのコンピュータおよびコンピュータネットワークには、生のデジタルビデオを処理するリソースが欠けている。この理由から、エンジニアは、圧縮（コーディングまたはエンコーディングとも称する）を使用して、デジタルビデオのビットレートを下げる。圧縮は、無損失とすることができ、この場合には、ビデオの品質は影響を受けず、ビットレートの減少が、ビデオの複雑さによって制限される。あるいは、圧縮を有損失とすることができ、この場合には、ビデオの品質に影響があるが、ビットレートの減少が、より劇的になる。圧縮解除では、圧縮の逆を行う。

【0004】一般に、ビデオ圧縮技法には、イントラフレーム圧縮（intraframe compression）とインターフレーム圧縮（interframe compression）が含まれる。イントラフレーム圧縮技法では、通常は1フレームまたはキーフレームと呼ばれる個々のフレームを圧縮する。インターフレーム圧縮技法では、通常は予測フレーム、Pフレーム、またはBフレームと呼ばれる、前のフレームおよび/または後のフレームを参照してフレームを圧縮する。

【0005】Microsoft Corporation社のWindows（登録商標）Media Video Version 8（「WMV8」）には、ビデオエンコーダおよびビデオデコーダが含まれる。WMV8エンコーダでは、イントラフレーム圧縮およびインターフレーム圧縮が使用され、WMV8デコーダでは、イントラフレーム圧縮解除およびインターフレーム

圧縮解除が使用される。WMV8エンコーダのインターフレーム圧縮では、ブロックに基づく動き補償された予測コーディングと、その後の、残留誤差の変換コーディングが使用される。

【0006】WMV8では、フレームが、3つのピクセルブレンすなわち、輝度ピクセル値の輝度（Y）ブレン、およびクロミナンスピクセル値の2つのクロミナンス（U、V）ブレンとして表される。Yブレンの解像度は、水平および垂直でUブレンおよびVブレンの解像度の2倍である。従って、320ピクセル×240ピクセルのフレームは、320ピクセル×240ピクセルのYブレンおよび160ピクセル×120ピクセルのUブレンおよびVブレンを有する。

【0007】WMV8エンコーダでは、予測フレームがピクセルの8×8ブロックに分割される。4つの8×8輝度ブロック、および2つの同一位置の8×8クロミナンスブロック（一方はUクロミナンスブレン、他方はVクロミナンスブレンに関する）のグループによって、16×16マクロブロックとが形成される。従って、16×16マクロブロックのそれぞれに、4つの8×8輝度ブロックと、2つの8×8クロミナンスブロックとが含まれる。

【0008】予測フレームのマクロブロックについて、WMV8エンコーダでは、動き予測が実行される。動き予測では、予測フレーム内でマクロブロックを検索し、基準フレームからのマクロブロックとマッチングすることによって、予測フレーム内のマクロブロックの動きを近似する。例えば、図1では、WMV8エンコーダによって、予測フレーム（110）内のマクロブロック（115）の動きベクトルが計算される。この動きベクトルを計算するために、エンコーダは、基準フレーム（130）の検索領域（135）内を検索する。検索領域（135）内で、エンコーダは、よい一致を見つけるために、予測フレーム（110）からのマクロブロック（115）の輝度値を、基準フレーム（130）からのさまざまな候補ブロックの輝度値と比較する。WMV8エンコーダは、動きベクトル精度を切り替えることができ、整数ピクセル、1/2ピクセル、または1/4ピクセルの水平分解能と、整数ピクセルまたは1/2ピクセルの垂直分解能とを有する検索範囲および動きベクトルを使用することができる。サブピクセル精度の動きベクトルを用いると、WMV8エンコーダによって、ビデオシーケンス内のサブピクセルの動きを近似することができる。

【0009】動き補償中に、WMV8エンコーダでは、予測フレームのマクロブロックの動きベクトルを使用して、基準フレームからマクロブロックの予測子を選択する。動き予測されるマクロブロックのそれぞれについて、WMV8エンコーダでは、元のマクロブロックとその予測子の間の差（残差または誤差とも称する）を計算す

る。WMV8エンコーダでは、残差をブロックに分割し、残差ブロックを有損失圧縮する。予測フレームの動き予測されたマクロブロックを再構成するために、WMV8エンコーダでは、残差を圧縮解除し、各々のマクロブロックの予測子に加算する。

【0010】WMV8デコーダでも、予測フレームのマクロブロックの動きベクトルを使用して、基準フレームからマクロブロックの予測子を判定する。予測フレームの動き予測されたマクロブロックを再構成するために、WMV8デコーダでは、残差を圧縮解除し、マクロブロッ

クの予測子に加算する。

【0011】動き予測中または動き補償中に、動きベクトルがサブピクセル精度を有する（すなわち、 $1/2$ ピクセル）

$$H_0 = (f + g + R_x) > 1 \quad (1)$$

$$H_1 = (f + j + R_x) > 1 \quad (2)$$

$$H_2 = (f + g + j + k + R_x) > 2 \quad (3)$$

ここで、 $R_x$  および  $R_y$  は、特定のフレームの丸めモードを示す1ビット丸め制御フラグによって制御される丸め制御値である。丸め制御フラグに0がセットされている場合には、 $R_x = 2$ かつ $R_y = 1$ である。丸め制御フラグに1がセットされている場合には、 $R_x = R_y = 0$ である。丸め制御フラグの値は、Pフレームごとに1と0の間で交互に変える。各1フレームでは、丸め制御フラグの値が0にリセットされる。従って、丸め制御は、フレーム単位で動作する。

【0013】式1、2、および3は、双一次補間 (bilinear interpolation) の例である。双一次補間は、高速であり、ピクセル値が滑らかな傾向がある。この平

$$H_0 = (-e + 9f + 9g - h + 8) > 4 \quad (4)$$

$$H_1 = (-b + 9f + 9j - n + 8) > 4 \quad (5)$$

$$H_2 = (-t_0 + 9t_1 + 9t_2 - t_3 + 8) > 4 \quad (6)$$

ここで、 $t_0$ 、 $t_1$ 、 $t_2$ 、 $t_3$  は、次のように計算される。

$$t_0 = (-a + 9b + 9c - d + 8) > 4 \quad (7)$$

$$t_1 = (-e + 9f + 9g - h + 8) > 4 \quad (8)$$

$$t_2 = (-i + 9j + 9k - l + 8) > 4 \quad (9)$$

$$t_3 = (-m + 9n + 9o - p + 8) > 4 \quad (10)$$

【0015】式(4)から(10)は、入力値の範囲外の出力をもたらすことができる。例えば、8ビット入力（範囲0、...、255）について、一連の値、0 255 255 0によって、式(4)から(10)のどれにおいても、287の出力値が作られる。従って、WMV8エンコーダまたはWMV8デコーダでは、有効な範囲内になるように、すべての式(4)から(10)の出力値をクランプ（または「クリッピング」）する。例えば、8ビット出力値について、0未満の値は0に変更され、255を超える値は255に変更される。クランプは、範囲の問題に対処するが、計算の速度が落ちる。さらに、クランプは、精度の消失をもたらす。

【0016】WMV8エンコーダまたはWMV8デコーダでは、その後、補間の後続ステージで、ある $1/4$ ピ

セルまたは $1/4$ ピクセル）ときに、WMV8エンコーダまたはWMV8デコーダでは、基準フレーム内のサブピクセル位置のピクセル値を判定しなければならない。WMV8エンコーダまたはWMV8デコーダでは、補間フィルタを使用して、サブピクセル位置の値を生成する。図2に、整数ピクセル値 $a$ 、 $b$ 、 $c$ 、...、 $p$ の補間によって計算された値を有するサブピクセルサブ位置 $H_0$ 、 $H_1$ 、 $H_2$ を示す。

【0012】 $1/2$ ピクセル動きベクトル精度を用いて動作するときに、3つの別個の $1/2$ ピクセル位置 $H_0$ 、 $H_1$ 、 $H_2$ の輝度ピクセル値について使用される補間フィルタは、次の通りである。

※滑化は、望ましい効果（量子化雑音の知覚可能性の低下など）を有する場合があるが、有効なピクセル情報の消失につながる可能性もある。

【0014】 $1/4$ ピクセル動きベクトル分解能の場合に、WMV8エンコーダまたはWMV8デコーダでは、まず、双三次フィルタ (bicubic filter) を使用して、 $1/2$ ピクセル位置での輝度ピクセル値を補間する。双三次補間は、双一次補間より低速であるが、エッジ値が保存される傾向があり、より少ない有効ピクセル情報の消失をもたらす。3つの別個の $1/2$ ピクセル位置 $H_0$ 、 $H_1$ 、 $H_2$ の双三次フィルタは、次の通りである。

$$H_0 = (-e + 9f + 9g - h + 8) > 4 \quad (4)$$

$$H_1 = (-b + 9f + 9j - n + 8) > 4 \quad (5)$$

$$H_2 = (-t_0 + 9t_1 + 9t_2 - t_3 + 8) > 4 \quad (6)$$

$$t_0 = (-a + 9b + 9c - d + 8) > 4 \quad (7)$$

$$t_1 = (-e + 9f + 9g - h + 8) > 4 \quad (8)$$

$$t_2 = (-i + 9j + 9k - l + 8) > 4 \quad (9)$$

$$t_3 = (-m + 9n + 9o - p + 8) > 4 \quad (10)$$

セル位置のピクセル値を計算する。これら $1/4$ ピクセル位置は、水平方向に、2つの $1/2$ ピクセル位置の間または整数ピクセル位置と $1/2$ ピクセル位置との間のいずれかに配置される。これら $1/4$ ピクセル位置に関して、WMV8エンコーダまたはWMV8デコーダでは、丸め制御なしで、2つの水平に隣接する $1/2$ ピクセル/整数ピクセル位置を使用する双一次補間（すなわち $(x+y+1)/2$ ）が使用される。

【0017】輝度動きベクトルを計算したならば、WMV8エンコーダまたはWMV8デコーダでは、同一位置のクロミナンス動きベクトルを導出する。WMV8のクロミナンスブレンは、水平と垂直の両方で輝度ブレンの半分の大きさなので、輝度動きベクトル値を、適当なクロミナンス動きベクトル値にスケーリングしなけれ

ばならない。WMV8では、この変換処理に、輝度動きベクトルを半分にすること、結果のクロミナンス動きベクトルを1/2ピクセル精度に丸めることが含まれる。従って、1/2ピクセル精度を有する輝度動きベクトルは、1/4ピクセル精度を有するクロミナンス動きベクトルに変換されない。さらに、WMV8でのクロミナンス丸めは、ユーザによる修正または選択が可能ではない単一のモードで動作する。

【0018】WMV8では、基準フレーム内のサブピクセル位置のピクセル値が、いくつかの状況でアンダーフローまたはオーバーフローを示す場合がある。例えば、1/4ピクセル位置の輝度ピクセル値は、隣接する整数ピクセル位置の値が255であり、隣接する1/2ピクセル位置の値が287(0+9×255+9×255-0+8>>4=287)である場合に、271

(0,...,255の範囲の外)になる可能性がある(255+287+1>>1=271)。この問題に対処するために、マクロブロックについて残差ブロックを予測子に加算した後に、WMV8エンコーダおよびWMV8デコーダでは、必要な場合に、範囲0,...,255内になるようにマクロブロックの再構成された値をクランプする。

【0019】WMV8の他に、複数の国際標準規格が、ビデオの圧縮および圧縮解除に関係する。これらの標準規格には、国際電気通信連合[ITU]のMotion Picture Experts Group[MPEG]1、2および4標準規格と、H.261標準規格、H.262標準規格、およびH.263標準規格が含まれる。WMV8と同様に、これらの標準規格では、イントラフレーム圧縮およびインターフレーム圧縮の組合せが使用されるが、これらの標準規格は、通常は、使用される圧縮技法の詳細においてWMV8と異なる。

【0020】複数の標準規格(例えばMPEG4およびH.263)が、双一次フィルタおよび基本的な丸め制御を使用する1/2ピクセル動き予測および1/2ピクセル動き補償を提供する。さらに、H.263では、理論的には1/4ピクセル分解能(すなわち、1/2ピクセル輝度動きベクトルの分解能の半分)を有するクロミナンス動きベクトルが、1/2ピクセル精度またはフルピクセル精度のいずれかに丸められ、その結果、1/4ピクセル値がクロミナンス空間で許容されなくなる。標準規格の動き予測/補償に関する詳細については、各標準規格の仕様書自体を参照されたい。

【0021】

【発明が解決しようとする課題】動き予測および動き補償は、効果的な圧縮技法であるが、さまざまな上述した動き予測/補償技法(WMV8および上述した標準規格)の2、下記を含む複数の短所を有する。

【0022】(1)基準フレーム内のサブピクセル位置でピクセル値を計算するときに、エンコーダおよびデ

コーダが、中間値の精度を不必要に失う。例えば、WMV8で1/4ピクセル位置のピクセル値を計算するときに、1/2ピクセル位置の中間値が、より多数のビット深さを使用可能であるという事実にもかかわらず、4ビットだけ右シフトされる。さらに、WMV8エンコーダ/デコーダでは、1/4ピクセル位置の2ステージ補間中に中間値がクランプされ、これによって、計算速度が低下し、精度の不必要な消失がもたらされる。

【0023】(2)1/4ピクセル動き予測および動き補償のピクセル値の補間が、多くの場合に非効率である。例えば、WMV8では、1次元1/4ピクセル位置の計算が、1/2ピクセル位置に関するフィルタの使用とその後の双一次フィルタの使用を必要とする。

【0024】(3)エンコーダおよびデコーダは、複数ステージ補間で作成される可能性のある丸め誤差の累積を考慮に入れることができない。丸め誤差は、例えば、ピクセル値が、ビデオシーケンスのフレームからフレームへと繰り返して丸められるときに発生する。この丸め誤差によって、低品質低ビットレートビデオシーケンスの知覚可能な誤差(artifact)が引き起こされる可能性がある。例えば、WMV8エンコーダおよびWMV8デコーダで、複数のステージで1/4ピクセル位置のピクセル値について補間するときに、丸め制御が使用されない。その代わりに、各ステージの結果が、補間の各ステージで同一の形で(丸め制御なしで)丸められる。

【0025】(4)クロミナンスの丸めは、1/4ピクセル精度で実行されず、クロミナンス動きベクトル丸めオプションに対する制御が与えられない。例えば、WMV8エンコーダおよびWMV8デコーダは、すべてのクロミナンス動きベクトルを1/2ピクセル値に丸め、単一モードでのみ動作する。

【0026】動き予測および動き補償のディジタルビデオに対するクリティカルな重要性を与えられれば、動き予測および動き補償がよく開発された分野であることは、驚くべきものではない。しかし、前記の動き予測技法および動き補償技法の利益がどれほどであれ、それらは、下記に記述した技法およびツールの長所を有しない。

【0027】

【課題を解決するための手段】要約すると、この詳細な説明は、動き推定および動き補償のさまざまな技法およびツールを対象とする。例えば、ビデオのコーディングおよびデコーディングの応用分野において、複数ステージの補間を使用して、特定のサブサンプル位置の値を計算することができる。第1ステージまたは中間ステージで計算される中間値は、最終的な値のダイナミックレンジより低いダイナミックレンジを有する。第1ステージまたは中間ステージのクランプを、スキップすることができる。第1ステージまたは中間ステージで実行されるビットシフトを、最終ステージまで延期することもできる。独立にまたは組み合わせで使用することができるさ

さまざまな技法およびツールによって、サブサンプル位置の最終的な値が計算される精度を改善することができる。この高められた精度によって、圧縮ビデオシーケンスの総合的な品質が改善される。

【0028】一態様では、ビデオエンコーダおよびビデオデコーダなどのコンポーネントは、1または複数の近似サブサンプル位置における1または複数の中間値の少なくとも一部に基づいて、特定のサブサンプル位置の最終的な値を計算する。最終的な値は、 $x$  ビットの最終的なダイナミックレンジを有し、1または複数の中間値は、 $y$  ビットの中間ダイナミックレンジを有し、 $y$  は  $x$  より大きい。

【0029】もう1つの態様では、ビデオエンコーダまたはビデオデコーダは、複数ステージの補間を使用して、複数のビデオフレームの基準ビデオフレームにおける小数ピクセル位置の最終的なピクセル値を計算する。複数のステージは、少なくとも第1ステージおよび最終ステージを含む。計算には、第1ステージの1または複数の中間ピクセル値のクランプをスキップすることを含む。計算は、さらに、最終ステージの基準ビデオフレーム内の小数ピクセル位置における最終的なピクセル値のクランプを含む。

【0030】もう1つの態様では、ビデオエンコーダまたはビデオデコーダなどのコンポーネントは、複数ステージの補間を使用して、特定の小数サンプル位置の値を計算する。計算には、値の精度を高めるために、複数ステージの第1ステージから複数ステージの第2ステージへのビットシフトの延期を含む。

【0031】この技法およびツールは、従来技術の動き予測および動き補償の複数の短所に対処する。さまざまな技法およびツールを、組み合わせまたは独立に使用することができる。追加の特徴および長所は、添付図面に関して進められる以下の詳細な説明から明白になる。

【0032】

【発明の実施の形態】 本明細書に記載の実施形態は、動き予測および動き補償におけるサブピクセル補間の技法およびツールに関する。さまざまな実施形態は、補間の後のステージまでクランプおよび/またはビットシフト（精度の消失をもたらす可能性がある動作）を延期することによって、複数ステージ補間で精度を保つ技法およびツールに関する。他の実施形態は、複数ステージ補間の効率的なフィルタリング動作または丸め動作に関する。

【0033】エンコーダまたはデコーダは、基準フレーム上で、1または複数のブロックまたはマクロブロックなどのフレームの一部分上で、サブピクセル補間を実行する。エンコーダ/デコーダは、基準フレーム内のサブピクセル位置のピクセル値を計算する。エンコーダ/デコーダは、その後、サブピクセル精度の動きベクトルを使用して、動き補償を実行することができる。

【0034】いくつかの実施形態で、ビデオエンコーダまたはビデオデコーダは、ビデオコーディングアプリケーションまたはビデオデコーディングアプリケーションにおいてサブピクセル補間を実行する。その代わりに、別のエンコーダまたはデコーダ、あるいは別のタイプのコンポーネントは、サブピクセル補間または以下で説明する別のタイプのアプリケーションで他の技法を実行する。

【0035】基準フレームに対してサブピクセル補間を実行する代わりに、いくつかの実施形態で、エンコーダ/デコーダは、フィールド、オブジェクトレイトラ、または他のイメージに対するサブピクセル補間を実行する。

【0036】いくつかの実施形態では、サブピクセル補間は、YUV色空間で基準フレームの輝度プレーンおよびクロミナンスプレーンのピクセル値を計算することによって行われる。代替案では、色空間が異なる（例えば、YIQまたはRGB）。

【0037】さまざまな技法およびツールは、組み合わせまたは独立に使用することができる。異なる実施形態によって、1または複数の本明細書に記載の技法およびツールが実施される。これらの技法の動作を、典型的に、提示のために特定のシーケンス順で説明するが、この説明の方法に、特に順序付けが必要でない限り、動作の順序の小さな再構成が含まれることを理解されたい。例えば、シーケンス順に説明される動作を、いくつかの場合に、再配置するか同時に実行することができる。さらに、説明を簡単にするために、流れ図に、通常は、特定の技法を他の技法と共に使用することができるさまざまな方法を図示しない。

【0038】いくつかの実施形態で、ビデオエンコーダおよびビデオデコーダは、ビットストリーム内のさまざまなフラグおよび信号を使用する。特定のフラグおよび信号を説明するが、この説明の方法に、フラグおよび信号に関する異なる規約（例えば1ではなく0）が含まれることを理解されたい。

【0039】（1. コンピューティング環境）図3に、本明細書に記載の複数の実施形態をその中で実施することができる、適切なコンピューティング環境（300）を示す。この技法およびツールを、異なる汎用または特定用途のコンピューティング環境で実施することができるので、コンピューティング環境（300）は、使用または機能性に関する制限を暗示することを意図したものではない。

【0040】図3を参照すると、コンピューティング環境（300）に、少なくとも1つの処理ユニット（310）およびメモリ（320）が含まれる。図3では、この最も基本的な構成（330）が、破線の中に含まれる。処理ユニット（310）は、コンピュータ実行可能命令を実行し、実際のプロセッサまたは仮想プロセッサとすることができる。マルチプロセッシングシステムで



は、複数の処理ユニットがコンピュータ実行可能命令を実行して、処理能力を高める。メモリ(320)は、揮発性メモリ(例えば、レジスタ、キャッシュ、RAM)、不揮発性メモリ(例えば、ROM、EEPROM、フラッシュメモリなど)、またはこの2つの組合せとすることができる。メモリ(320)には、ビデオエンコーダおよび/またはビデオデコーダなどのエンコーダおよび/またはデコーダでサブピクセル補間技法を実施するソフトウェア(380)が格納される。

[0041] コンピューティング環境は、追加の特徴を有する場合がある。例えば、コンピューティング環境(300)に、ストレージ(340)、1または複数の入力デバイス(350)、1または複数の出力デバイス(360)、および1または複数の通信接続(370)が含まれる。バス、コントローラ、またはネットワークなどの相互接続機構(図示せず)によって、コンピューティング環境(300)のコンポーネントが相互接続される。通常、オペレーティングシステムソフトウェア(図示せず)によって、コンピューティング環境(300)内で実行される他のソフトウェアのオペレーティング環境が提供され、コンピューティング環境(300)のコンポーネントのアクティビティが調整される。

[0042] ストレージ(340)は、取外し可能または取外し不能とすることができ、ストレージ(340)には、磁気ディスク、磁気テープ、磁気カセット、CD-ROM、DVD、または、情報を格納できコンピューティング環境(300)内でアクセスできる他の任意の媒体が含まれる。ストレージ(340)には、サブピクセル補間技法を実施するソフトウェア(380)の命令が格納される。

[0043] 入力デバイス(350)は、キーボード、マウス、ペン、またはトラックボールなどの接触入力デバイス、音声入力デバイス、スキャニングデバイス、または、コンピューティング環境(300)に入力を提供する別のデバイスとすることができる。オーディオエンコードまたはビデオエンコードのために、入力デバイス(350)を、サウンドカード、ビデオカード、TVチューナカード、またはアナログ形式またはデジタル形式でオーディオ入力またはビデオ入力を受け入れる同様のデバイス、あるいは、オーディオサンプルまたはビデオサンプルをコンピューティング環境(300)に読み込むCD-ROMまたはCD-RWとすることができる。出力デバイス(360)は、ディスプレイ、プリンタ、スピーカ、CDライター、または、コンピューティング環境(300)からの出力を提供する別のデバイスとすることができる。

[0044] 通信接続(370)によって、通信媒体を介する別のコンピューティングエンティティへの通信が可能になる。通信媒体は、コンピュータ実行可能命令、オーディオまたはビデオの入力または出力、あるいは他

のデータなどの情報が、変調されたデータ信号で伝えられる。変調されたデータ信号とは、信号内で情報をエンコードする方法により、1または複数の特性を設定されまたは変更された信号である。限定ではなく例として、通信媒体には、電気、光、RF、赤外線、音響、または他の搬送波を用いて実施される有線もしくは無線の技術が含まれる。

[0045] 技法およびツールを、コンピュータ読み取り可能媒体の全般的な文脈で説明することができる。コンピュータ読み取り可能媒体とは、コンピューティング環境内でアクセスできるすべての使用可能な媒体である。制限ではなく例として、コンピューティング環境(300)に関して、コンピュータ読み取り可能媒体に、メモリ(320)、ストレージ(340)、通信媒体、およびこれらの任意の組合せが含まれる。

[0046] 技法およびツールを、プログラムモジュールに含まれるものなど、目的とする現実のプロセスまたは仮想プロセス上のコンピューティング環境内で実行されるコンピュータ実行可能命令の全般的な文脈で説明することができる。一般に、プログラムモジュールには、特定のタスクを実行するか特定の抽象データ型を実施する、ルーチン、プログラム、ライブラリ、オブジェクト、クラス、コンポーネント、データ構造などが含まれる。望みに応じてさまざまな実施形態において、プログラムモジュールの機能性を組み合わせるか、プログラムモジュールの間で分割することができる。プログラムモジュールのコンピュータ実行可能命令を、ローカルコンピューティング環境または分散コンピューティング環境内で実行することができる。

[0047] 提示のために、この詳細な説明で、「決定」および「選択」などの用語を使用して、コンピューティング環境でのコンピュータ動作を説明する。これらの用語は、コンピュータによって実行される動作の高水準の抽象化であり、人間によって実行される動作と混同してはならない。これらの用語に対応する実際のコンピュータ動作は、実施形態に応じて変化する。

[0048] (11. 一般化されたビデオエンコーダおよびビデオデコーダ) 図4は、一般化されたビデオエンコーダ(400)のブロック図であり、図5は、一般化されたビデオデコーダ(500)のブロック図である。

[0049] エンコーダおよびデコーダ内のモジュールの間に示された関係は、エンコーダおよびデコーダ内の情報の主な流れを示し、図を簡単にするために、他の関係は図示されていない。具体的に言うと、図4および図5には、通常は、ビデオシーケンス、フレーム、マクロブロック、ブロックなどに使用されるエンコーダ設定、モード、テーブルなどを示すサイド情報(side information)が示されていない。そのようなサイド情報は、通常はサイド情報のエントロピー符号化(entropy encoding)の後に、出力ビットストリーム内で送達される。出

力ビストリートのフォーマットは、Windows（登録商標）Media Videoフォーマットまたは別のフォーマットとすることができ。

【0050】エンコーダ（400）およびデコーダ（500）は、ブロックベースであり、4:2:0マクロブロックフォーマットを使用し、各マクロブロックには、4つの8×8輝度ブロック（時には1つの16×16マクロブロックとして扱われる）と、2つの8×8クロミナンスブロック（例えば、1つはUブロック、1つはVブロック）が含まれる。代替案では、エンコーダ（400）およびデコーダ（500）が、オブジェクトベースであり、異なるマクロブロックフォーマットまたはブロックフォーマットを使用するか、8×8ブロックおよび16×16マクロブロックと異なるサイズまたは構成のピクセルの組に対する操作を実行する。

【0051】実施形態および所望の圧縮のタイプに応じて、エンコーダまたはデコーダのモジュールを、追加し、省略し、複数のモジュールに分割し、他のモジュールと組み合わせ、かつ/または類似するモジュールで置換することができる。代替実施形態では、異なるモジュールおよび/またはモジュールの他の構成を有するエンコーダまたはデコーダによって、本明細書に記載の技法の1または複数が実行される。

【0052】(A. ビデオエンコーダ) 図4は、一般的なビデオエンコーダシステム（400）のブロック図である。エンコーダシステム（400）は、現在のフレーム（405）を含むビデオフレームのシーケンスを受け取り、出力として圧縮ビデオ情報（495）を作る。ビデオエンコーダの特定の実施形態では、通常は、一般化されたエンコーダ（400）の変形形態または補足されたバージョンが使用される。

【0053】エンコーダシステム（400）によって、予測フレームおよびキーフレームが圧縮される。提示のために、図4に、エンコーダシステム（400）を介するキーフレームのパスと、順方向予測フレームのパスを示す。エンコーダシステム（400）のコンポーネントの多くが、キーフレームと予測フレームの両方の圧縮に使用される。これらのコンポーネントによって実行される正確な動作を、圧縮される情報のタイプに応じて変更することができる。

【0054】予測フレーム〔インターコーディングされたフレーム、あるいは両方向予測の場合にbフレームまたはbフレームとも称する〕は、あるフレームから他のフレームへの予測（または差）に関して表現される。予測残差は、予測されたものと元のフレームとの差である。対照的に、キーフレーム〔iフレーム、イントラコーディングされたフレームとも称する〕は、他のフレームへの参照なしで圧縮される。

【0055】現在のフレーム（405）が、順方向予測フレームである場合には、動き予測器（410）が、基

準フレームに関して、現在のフレーム（405）のマクロブロックまたはピクセルの他の組の動きを推定し、基準フレームは、フレームストア（420）にバッファリングされ、再構成された前のフレーム（425）である。代替実施形態では、基準フレームが、後のフレームであるか、現在のフレームであるか、両方向予測される。動き予測器（410）は、サイド情報として、動きベクトルなどの動き情報（415）を出力する。動き補償器（430）が、再構成された前のフレーム（425）に動き情報（415）を適用して、動き補償された現在のフレーム（435）を形成する。しかし、予測は、ほとんどの場合に完全ではなく、動き補償された現在のフレーム（435）と元の現在のフレーム（405）との差が、予測残差（445）である。代替案では、動き予測器および動き補償器が、別のタイプの動き予測/補償を適用する。

【0056】周波数変換器（460）が、空間領域ビデオ情報を周波数領域（すなわちスペクトル）データに変換する。ブロックベースのビデオフレームについて、周波数変換器（460）は、離散コサイン変換〔DCT〕またはDCTの変形形態を動き予測残差データのブロックに適用して、DCT係数のブロックを作る。代替案では、周波数変換器（460）が、フーリエ変換などの従来の周波数変換を適用するか、ウェーブレットまたはサブバンド（subband）分析を使用する。いくつかの実施形態で、周波数変換器（460）が、キーフレームの空間予測残差のブロックに周波数変換を適用する。周波数変換器（460）は、8×8、8×4、4×8、または他のサイズの周波数変換を適用することができる。

【0057】その後、量子化器（470）が、スペクトルデータ係数のブロックを量子化する。量子化器は、フレームごとまたは別の基準で変化するステップサイズを用いてスペクトルデータに均一の量子化を適用する。代替案では、量子化器が、例えば不均一量子化、ベクトル量子化、または非適応量子化などの別のタイプの量子化をスペクトルデータ係数に適用し、あるいは、周波数変換を使用しないエンコーダシステムでは、スペクトル領域データを直接に量子化する。適応量子化のほかに、エンコーダ（400）は、フロッピング、適応フォールタリング、または他のレート制御の技法を使用することができる。

【0058】再構成された現在のフレームが、後続の動き予測/補償に必要なときには、逆量子化器（476）が、量子化されたスペクトルデータ係数に対して逆量子化を実行する。その後、逆周波数変換器（466）が、周波数変換器（460）の動作の逆を実行し、再構成された予測残差（予測フレームの場合）または再構成されたキーフレームを作る。現在のフレーム（405）が、キーフレームである場合には、再構成されたキーフレーム

17

ムが、再構成された現在のフレーム（図示せず）として採用される。現在のフレーム（405）が、予測フレームである場合には、再構成された予測残差を動き補償された現在のフレーム（435）に加算して、再構成された現在のフレームを形成する。フレームストア（420）は、次のフレームの予測に使用するために、再構成された現在のフレームをバッファリングする。いくつかの実施形態で、エンコーダが、再構成されたフレームにデブロッキングフィルタ（deblocking filter）を適用して、フレームのブロック内の不連続性を適応的に平滑化する。

【0059】エントロピ符号化器（480）は、量子化器（470）の出力ならびにあるサイド情報（例えば、動き情報（415）、量子化ステップサイズなど）を圧縮する。通常のエントロピ符号化技法には、算術コーディング、差分コーディング、ハフマンコーディング、ランレンスコーディング、LZコーディング、辞書コーディング、および上記の組合せが含まれる。エントロピ符号化器（480）は、通常は、異なる種類の情報（例えば、DC係数、AC係数、異なる種類のサイド情報）に異なる符号化技法を使用し、特定の符号化技法内で複数のコードテーブルの中から選択することができる。

【0060】エントロピ符号化器（480）は、圧縮ビデオ情報（495）をバッファ（490）に入れる。バッファレベルインジケータが、ビットレート適応モジュールにフィードバックされる。圧縮ビデオ情報（495）は、一定のまたは比較的に一定のビットレートでバッファ（490）から放出され、そのビットレートでの後続のストリーミングのために格納される。代替案では、エンコーダシステム（400）が、圧縮の直後に圧縮ビデオ情報をストリーミングする。

【0061】バッファ（490）の前または後に、圧縮ビデオ情報（495）を、ネットワークを介する伝送のためにチャネルコーディングすることができる。チャネルコーディングでは、エラー検出および訂正データを圧縮ビデオ情報（495）に適用することができる。

【0062】（B. ビデオデコーダ）図5は、一般的なビデオデコーダシステム（500）のブロック図である。デコーダシステム（500）は、ビデオフレームの圧縮されたシーケンスに関する情報（595）を受け取り、再構成されたフレーム（505）を含む出力を作る。ビデオデコーダの特定の实施形態では、通常は、一般化されたデコーダ（500）の変形形態または補足された版が使用される。

【0063】デコーダシステム（500）は、予測フレームおよびキーフレームを圧縮解除する。提示のために、図5に、デコーダシステム（500）を介するキーフレームのパスおよび順方向予測フレームのパスを示す。デコーダシステム（500）のコンポーネントの多くが、キーフレームおよび予測フレームの両方の圧縮解

18

除に使用される。これらのコンポーネントによって実行される正確な動作を、圧縮解除される情報のタイプに応じて変更することができる。

【0064】バッファ（590）が、圧縮ビデオシーケンスに関する情報（595）を受け取り、受け取った情報をエントロピ復号化器（580）から使用可能にする。バッファ（590）は、通常は、経時的にかなり一定の速度で情報を受け取り、バッファ（590）には、帯域幅または伝送の短期間変動を平滑化するためにジッタバッファが含まれる。バッファ（590）に、再生バッファおよび他のバッファも含めることができる。代替案では、バッファ（590）が、変化する速度で情報を受け取る。バッファ（590）の前または後に、圧縮ビデオ情報を、チャネルデコードし、エラー検出および訂正のために処理することができる。

【0065】エントロピ復号化器（580）は、通常はエンコーダ内で実行されるエントロピ符号化の逆を適用することによって、エントロピ符号化された量子化されたデータならびにエントロピ符号化されたサイド情報（例えば、動き情報（515）、量子化ステップサイズ）をデコードする。エントロピ復号化技法には、算術コーディング、差分デコーディング、ハフマンデコーディング、ランレンスデコーディング、LZデコーディング、辞書デコーディング、および上記の組合せが含まれる。エントロピ復号化器（580）は、頻繁に、異なる種類の情報（例えば、DC係数、AC係数、異なる種類のサイド情報）に異なる復号化技法を使用し、特定の復号化技法内で複数のコードテーブルの中から選択することができる。

【0066】再構成されるフレーム（505）が、順方向予測フレームである場合には、動き補償器（530）が、動き情報（515）を基準フレーム（525）に適用して、再構成されるフレーム（505）の予測（535）を形成する。例えば、動き補償器（530）は、マクロブロック動きベクトルを使用して、基準フレーム（525）内のマクロブロックを見つける。フレームバッファ（520）に、基準フレームとして使用される、前に再構成されたフレームが格納される。代替案では、動き補償器が、別のタイプの動き補償を適用する。動き補償器による予測は、ほとんどの場合に完全ではなく、従って、デコーダ（500）は、予測残差も再構成する。

【0067】デコーダが、後続の動き補償のために、再構成されたフレームを必要とするときに、フレームストア（520）に、再構成されたフレームが、次のフレームの予測に使用するために格納される。いくつかの実施形態で、エンコーダは、再構成されたフレームにデブロッキングフィルタを適用して、フレームのブロック内の不連続性を適応的に平滑化する。

【0068】逆量子化器（570）が、エントロピ復号

50

化されたデータを逆量子化する。一般に、逆量子化器は、フレームごとまたは別の基準で変化するステップサイズを用いてエントロピー符号化されたデータに均一のスカラー逆量子化を適用する。代替案では、逆量子化器が、例えば不均一逆量子化、ベクトル逆量子化、または非適応逆量子化など別のタイプの逆量子化をデータに適用し、または、逆周波数変換を使用しないデコーダシステムでは、空間領域データを直接に逆量子化する。

【0069】逆周波数変換器(560)は、量子化された周波数領域データをスペクトル領域ビデオ情報に変換する。ブロックベースビデオフレームについて、逆周波数変換器(560)は、逆DCT[「IDCT」]またはIDCTの変形をDCT係数のブロックに適用し、動き予測残差データを作る。代替案では、逆周波数変換器(560)が、逆フーリエ変換など別の従来の逆周波数変換を適用するか、ウェーブレット分析またはサブバンド分析を使用する。いくつかの実施形態で、逆周波数変換器(560)は、キーフレームの空間予測残差のブロックに逆周波数変換を適用する。逆周波数変換器(560)は、 $8 \times 8$ 、 $8 \times 4$ 、 $4 \times 8$ 、または他のサイズの逆周波数変換を適用することができ、

【0070】(111、動き予測および動き補償)インターフレームコーディングでは、フレームの間の時間的冗長性を活用して、圧縮を達成する。時間的冗長性削減では、現在のフレームを符号化するときに、予測子として、前に符号化されたフレームを使用する。以下で説明する実施形態では、ビデオエンコーダで、より少ないビット数を使用して情報を符号化するために、通常のビデオシーケンス内の時間的冗長性を活用する。ビデオエンコーダでは、基準フレーム(例えば、前に符号化された、前のフレーム)に対する予測フレームのピクセルのブロック、マクロブロック、または他の組の動きをパラメータ化するのに動き予測を使用する。ビデオエンコーダ(ならびに対応するデコーダ)では、動き情報および基準フレームを使用して予測フレームを再構成するのに、動き補償を使用する。

【0071】動き補償は、基準フレームを変位させることによってビデオフレームの予測(すなわち、予測フレーム)を生成する処理である。上述したように、予測は、基準フレームからのデータのブロック、マクロブロック、または他の組について形成される。また、通常は、変位が、直線をなし、予測されるタイル全体にわたって一定である。そのような変位は、X方向およびY方向に沿った変位またはシフトに対応する2つの成分を有する動きベクトルによって定義される。X(水平)およびY(垂直)の動きベクトル成分は、現在予測された7つあるタイルと、基準フレーム内の対応する位置との間の変位を表す。正の値は、現在の位置の下および右の位置を表す。負の値は、現在の位置の上および左の位置を表す。

【0072】一実施形態では、ブロックが、ピクセルの $8 \times 8$ タイルであり、マクロブロックが、ピクセルの $16 \times 16$ タイルであり、動きベクトルが、 $1/4$ ピクセル精度で定義される。他の実施形態では、エンコーダおよびデコーダで、異なる分解能または任意の変化する動きベクトルを用いて、および/または動きベクトル以外の動き情報を使用して、本明細書に記載の技法の1または複数の、異なるサイズのタイルまたは任意の変化するサイズのタイルに適用する。

【0073】動きベクトル補償は、通常は、ピクセル変位に関して、しばしばサブピクセル精度を用いて、指定される。サブピクセル変位は、適当に定義された動き補償フィルタを使用して基準フレームをフィルタリングすることによって実現される。直線をなすサブピクセル動き補償の場合に、X成分およびY成分が、固定小数点の数として表現される。これらの数の整数部分を、フルピクセルシフトと称し、仮数部分を、サブピクセルシフトと称する。サブピクセルシフトが0のときに、動きは、整数値のピクセルである。しばしば、これは、予測子を生成するための基準フレームからのブロックコピーとして実施される(理論上は、何らかの形のフィルタリングを潜在的に適用できるはずであるが)。その一方で、サブピクセルシフトが0でないときには、サブピクセルシフトに対応する1または複数のフィルタを基準フレームの整数ピクセル位置に適用することによって、予測子が生成される。従って、動き補償フィルタは、サブピクセルシフトによって決定される。

【0074】フィルタリング動作としてサブピクセルシフトを実施するために、動き補償フィルタで、整数ピクセル位置の基準値に基づいて、小数ピクセル位置でデータ点を補間する。一般に、補間の品質は、フィルタのサポートに依って増加する。いくつかの実施形態では、分離可能な2タップおよび4タップ(各方向で)のフィルタが使用され、これらのフィルタは、二次インターポレータおよび三次インターポレータに対応する。

【0075】いくつかの実施形態で、動き補償フィルタが、整数算術と、ビットシフトとして実施される除算とを使用する。丸め制御パラメータRは、0または1の値をとり、これによって、これらの除算の丸めの方向が決定される。この丸め制御パラメータは、定数をセットされる。外部からシグナリングされ、または過去の符号化された情報から暗黙のうちに導出することができる。

【0076】図6に、いくつかの実施形態でサブピクセル動き予測およびサブピクセル動き補償中に使用される基準フレーム(600)内の整数ピクセル位置およびサブピクセル位置を示す。各方向の $1/4$ ピクセル間隔で、基準フレーム(600)に、エンコーダまたはデコーダが特定の変位についてピクセル値を補間する可能性があるサブピクセル位置が含まれる。基準フレーム(600)の整数位置aからpは、図6では影付きの円とし

て示され、整数位置の間で補間される1/4位置および \*セル位置を表す。

1/2位置は、影なしの円として示されている。位置P 【0077】

。からP<sub>9</sub>は、表1に記載の、9個の代表的なサブピクセル 【表1】

表1：代表的なサブピクセル位置

位置	説明
P <sub>0</sub>	水平に1/4ピクセル、垂直にフルピクセル
P <sub>1</sub>	水平に1/2ピクセル、垂直にフルピクセル
P <sub>2</sub>	水平にフルピクセル、垂直に1/4ピクセル
P <sub>3</sub>	水平に1/4ピクセル、垂直に1/4ピクセル
P <sub>4</sub>	水平に1/2ピクセル、垂直に1/4ピクセル
P <sub>5</sub>	水平にフルピクセル、垂直に1/2ピクセル
P <sub>6</sub>	水平に1/4ピクセル、垂直に1/2ピクセル
P <sub>7</sub>	水平に1/2ピクセル、垂直に1/2ピクセル
P <sub>8</sub>	水平にフルピクセル、垂直に3/4ピクセル

【0078】P<sub>8</sub>で例示される3/4ピクセル位置は、1/4ピクセル位置の特別なケースとみなすことができ、これは、フルピクセル位置から1/4ピクセルだけシフトされている。他の3/4ピクセル位置は、可能であるが図示されていない。サブピクセル位置P<sub>0</sub>からP<sub>8</sub>は、後の補間フィルタの説明で参照する。代替実施形態では、エンコーダおよびデコーダが、追加のまたは異なるサブピクセル位置、例えば各方向に1/4ピクセル以外の間隔で、値を補間する。

【0079】(A) 近似双三次補間フィルタ)いくつかの実施形態のサブピクセル補間について、ビデオエンコーダおよびビデオデコーダで、下記のように定義される一次/双三次フィルタおよび/または三次/双三次フィルタが使用される。

【0080】一次インターポーレータは、補間される点に最も近い2つの格子点で既知の値を使用する、線形または一次の一次元多項式である。補間される点での線形関数の値が、線形補間である。線形多項式の係数は、式の線形系を解き、線形フィルタの係数を決定することによって計算される。線形補間フィルタは、2つのフィルタタップによって定義される。双一次インターポーレータは、2つの次元で分離可能な線形インターポーレータである。

【0081】三次インターポーレータは、補間される点に最も近い4つの格子点での既知の値を使用する、立方または三次元多項式である。補間される点での三次関数の値が、三次補間である。三次多項式の係数は、式の系を解き、三次フィルタの係数を決定することによって計算される。三次インターポーレータフィルタは、4つの※

$$1/2 \text{ピクセルシフト } F_1: [-1 \ 9 \ 9 \ -1] \quad (11)$$

$$1/4 \text{ピクセルシフト } F_2: [-4 \ 53 \ 18 \ -3] \quad (12)$$

$$3/4 \text{ピクセルシフト } F_3: [-3 \ 18 \ 53 \ -4] \quad (13)$$

【0085】実際には、フィルタに、潜在的にフィルタ係数によって導入される拡大を補償するために、右シフトが含まれる(例えば、F<sub>1</sub>では4ビット、F<sub>2</sub>およびF<sub>3</sub>では6ビット)。演算子>>は、右シフト演算子で

※フィルタタップによって定義される。双三次インターポーレータは、2つの次元で分離可能な三次インターポーレータである。

【0082】一次および双一次という用語は、通常は、ビデオ圧縮およびビデオ圧縮解除の分野では交換可能に使用される。普通の2次元補間では、1次元で実行される補間動作が、他の次元に複製され、従って、各フィルタリングステージを、双一次フィルタリングと称する。三次および双三次という用語は、同様に交換可能である。

【0083】本明細書では、一次および双一次という用語は、1次元、2次元、または3次元以上のフィルタリングを説明するのに交換可能に使用される。同様に、三次および双三次という用語は、1次元、2次元、または3次元以上のフィルタリングを説明するのに交換可能に使用される。例えば、式(11)から(13)では、三次フィルタのタイプが定義されるが、これらは、双三次フィルタと呼ばれる。というのは、基準ビデオフレームの2ステージ補間の一般的な応用例で、フィルタが、2ステージ補間の両方の次元について複製される動作で使用されるからである。より一般的には、フィルタリングの次元数は、文脈から既知である。

【0084】いくつかの実施形態で、エンコーダおよびデコーダで、近似双三次フィルタを使用して、サブピクセル位置の値を補間する。例えば、エンコーダおよびデコーダで、図6に示されたものなどの基準フレームの可能なシフト位置で、下記のフィルタ(F<sub>1</sub>が双三次フィルタ、F<sub>2</sub>およびF<sub>3</sub>が近似双三次フィルタ)が使用される。

ある。右シフト演算子によって、2進数のビットが右にシフトされ、最下位ビットが捨てられ、最上位ビットに0が追加される。この演算は、剰余が切り捨てられる、シフトされるビット数の2のべきによる単純な除算をも

たらず(例えば、3による右シフトは、 $2^3=8$ による除算をもたらす)。

【0086】 $F_2$ および $F_3$ のフィルタ係数は、真の1/4ピクセル双三次インターポーレータ(4タップフィルタ)

$$(-7e + 105f + 35g - 5h) >> 7 \quad (14)$$

【0087】係数の値の合計が、128になり、フィルタリングの産物は、7ビットだけ右シフトされる。近似双三次フィルタ $F_2$ および $F_3$ は、性能に関して純粋な※

$$\begin{aligned} & (-7e + 105f + 35g - 5h) >> 7 \\ & = (-3.5e + 5.25f + 1.75g - 0.625h) >> 6 \\ & \approx (-4e + 5.3f + 1.8g - 0.6h) >> 6 \end{aligned} \quad (15)$$

【0088】多くの場合に、純粋な双三次フィルタを使用することによって、複数ステージの補間における精度のビット消失がもたらされ、従って、近似双三次フィルタに関する正規化係数が、少なくとも1/2だけ減らされる(すなわち、右シフトが1ビット以上減らされる)。式(15)の近似双三次フィルタについて選択されたフィルタ係数は、周波数領域の挙動(例えば、高周波数情報を保存するため)および経験的挙動(例えば、あるビットレートに対する最小のひずみを達成するため)を考慮に入れた後の、真の双三次フィルタの丸めに基づく。具体的に言うと、フィルタ $F_2$ および $F_3$ には、4つのフィルタ係数が含まれる(一般に、フィルタで使用されるフィルタ係数が少ないほど、実施が高速になるが、近接ピクセルの雑音に対処するために、十分なフィルタ係数を使用しなければならない)。フィルタ係数値は、合計が64になるように調節され、これによって、より高い分解能の双三次フィルタを近似しながら、16ビット算術を使用する実施形態が容易になる。双三次フィルタを近似しながら、合計が64になる他のフィルタ係数値を使用することもできる。実質的に純粋な双三次フィルタのように実行されるが、より少ないサポータおよび/またはより低い分解能を有するフィルタを、

「近似」双三次フィルタと称する。フィルタが実質的に純粋な双三次フィルタのように実行されるかどうかを客観的に測定する方法の1つが、近似フィルタが純粋な双三次フィルタによく相関する(すなわち、定義された閾値以内である)かをチェックすることである。一実施形態では、相関が、フィルタのベクトル間の角度のコサイン(できる限り1に近いことが望まれる)を測定することであり、閾値は0.95である。他の客観的または主★

$$F_1: (-1I_1 + 9I_2 + 9I_3 - 1I_4 + 8-r) >> 4 \quad (16)$$

$$F_2: (-4I_1 + 53I_2 + 18I_3 - 3I_4 + 32-r) >> 6 \quad (17)$$

$$F_3: (-3I_1 + 18I_2 + 53I_3 - 4I_4 + 32-r) >> 6 \quad (18)$$

ここで、値 $r$ によって、丸めが制御される。下記のように、値 $r$ は、2進フレームレベル丸め制御パラメータ $R$ と補間方向に依存する。 ☆

$$r = \begin{cases} 1-R & (\text{垂直方向}) \\ R & (\text{水平方向}) \end{cases} \quad (19)$$

【0093】1次元補間をさらに示すために、図6のP および $P_2$ によって、1つの次元だけでの補間を必要

※ルタ)に大雑把に基づく。下記の式に、位置 $P$ について真の1/4ピクセル双三次フィルタを適用した結果を示す。

★観的な測定、他の相関測定値、および/または閾値を、使用することもできる。例えば、近似双三次フィルタのフィルタ係数を選択し、その結果、それらの合計が、効率的なフーリエ変換または他の数学的演算を容易にする他の値になるようにすることができる。

【0089】以下でより完全に説明するように、図7に、式(11)から(13)で概要を示された双三次フィルタに対応する場合のそれぞれの補間されたピクセルを計算するのに使用されるピクセル値と共に、整数ピクセル位置を示す。 $P$ は、ピクセル値が計算されるサブピクセル位置を示す。 $I_1$ 、 $I_2$ 、 $I_3$ 、および $I_4$ は、補間の次元に沿った整数ピクセル位置を表す。図7には、水平補間が示されているが、同一の演算および位置の配置が、垂直補間に適用される。

【0090】代替実施形態では、エンコーダおよびデコーダは、他のおよび/または追加の補間フィルタを使用する。例えば、エンコーダおよびデコーダは、双一次(すなわち2タップの)フィルタを、値の補間に使用する。例えば、図6のサブピクセル位置を参照すると、 $P_1$ 、 $P_2$ 、および $P_3$ の値を決定するのに使用される補間フィルタを、式(1)から(3)に示されたフィルタとすることができる。

【0091】(B、1次元補間)さまざまなサブピクセル位置について、いくつかの実施形態のエンコーダおよびデコーダでは、補間された値を1つの次元だけで計算する。図7に示されているように、下記の式によって、整数ピクセルの間で補間するときの、フィルタ $F_1$ (1/2ピクセルシフト)、 $F_2$ (1/4ピクセルシフト)および $F_3$ (3/4ピクセルシフト)の動作が示される。

☆【0092】

【数1】

とする(すなわち、 $P_1$ では水平方向、 $P_2$ では垂直方向)、基準フレーム(600)内の1/2ピクセル位置が示されている。次式によって、 $P_1$ および $P_2$ について\*

$$P_1 = (-1e + 9f + 9g - 1h + 8 - r) >> 4 \quad (20)$$

$$P_2 = (-1b + 9f + 9j - 1n + 8 - r) >> 4 \quad (21)$$

【0094】同様に、図6の $P_2$ および $P_3$ によって、1つの次元だけでの補間を必要とする、基準フレーム(600)内の1/4ピクセル位置が示される。次式に※

$$P_3 = (-4e + 53f + 18g - 3h + 32 - r) >> 6 \quad (22)$$

$$P_4 = (-4b + 53f + 18j - 3n + 32 - r) >> 6 \quad (23)$$

近似1/4ピクセル双三次フィルタ $F_2$ を、わずかな修正だけを用いて使用して、3/4ピクセル位置を計算することもできる。例えば、次式によって、 $P_4$ について★

$$P_4 = (-3b + 18f + 53j - 4n + 32 - r) >> 6 \quad (24)$$

【0095】代替案では、エンコーダおよびデコーダで、1つの次元の1/2ピクセル、1/4ピクセル、または3/4ピクセルだけシフトされる位置について他のおよび/または追加の補間フィルタが使用される。例えば、エンコーダおよびデコーダで、より多数またはより少数のフィルタ係数、異なるフィルタ係数、異なる丸め、または丸めなしのフィルタが使用される。

【0096】(C. 多次元補間)いくつかの実施形態では、補間が、2次元でオフセットしたサブピクセル位置で実行される。例えば、図8で、 $P_3$ 、 $P_4$ 、 $P_5$ 、および $P_7$ が、水平と垂直の両方の次元で補間が行われる位置である。

【0097】図8に示された補間方法(800)に対応する一実施形態では、2次元サブピクセル位置が、まず垂直方向に沿って、次に水平方向に沿って、補間される。以下でより完全に説明するように、補間は、上の式(16)から(18)で指定されたフィルタ、 $F_1$ 、 $F_2$ 、または $F_3$ の1または複数を使用して実行される。図8に示された実施形態では、丸めが、垂直フィルタリングの後に水平フィルタリングの後の両方で適用される。丸め規則のビットシフトによって、中間結果での16ビット算術によって許容される精度の維持が保証される。

【0098】図8では、垂直フィルタリングがまず実行☆

$$(S + R_v) >> \text{shiftV}$$

ここで、 $S$ は、垂直にフィルタリングされた結果であり、 $R_v = 2^{\text{shiftV} - 1} + R$ である。 $R$ は、フレームごと0と1の間で交互に変わる丸め制御値である。従って、丸め規則には、ステージを交互に変わる(stage-alternating)丸め制御(813)およびビットシフト(814)が含まれる。

【0102】右シフトによって、潜在的に分解能の消失が引き起こされ、従って、右シフトの少なくとも一部が、補間の後のステージまで延期される。 $\text{shiftV}$ の右シフト値は、補間されるサブピクセル位置に依存する。具体的に言うと、 $P_3$ 、 $P_4$ 、 $P_5$ 、および $P_7$ に

※で整数ピクセルの間で補間するときのフィルタ $F_1$ (1/2ピクセルシフト)の動作が示される。

※によって、 $P_3$ および $P_4$ について整数ピクセルの間で補間するときのフィルタ $F_2$ (1/4ピクセルシフト)の動作が示される。

★整数ピクセルの間で補間するときのフィルタ $F_3$ (3/4ピクセルシフト)の動作が示される。

☆され、水平フィルタリングがそれに続く。垂直フィルタリングから開始することによって、いくつかのアーキテクチャで性能が改善される。他の実施形態では、フィルタリングの順序が異なる。例えば、補間が、垂直方向の前に水平方向で実行される。あるいは、補間フィルタのさまざまな他の組合せが使用される(例えば、複数の水平フィルタおよび/または複数の垂直フィルタ)。

【0099】入力ピクセル値(811)および出力ピクセル値(838)は、8ビットのビット深さを有し、256値のダイナミックレンジを有する。中間値(820)は、16ビットのビット深さを有し、65536値のダイナミックレンジを有する。代替実施形態では、入力値、出力値、および中間値が、異なる(例えばより大きい)ビット深さを有する。

【0100】第1ステージ(810)で、適当な垂直フィルタ( $F_v$ )が、8ビット入力ピクセル値(811)に適用される(812)。適用される垂直フィルタは、選択されたサブピクセル位置が、1/4ピクセル、1/2ピクセル、または3/4ピクセルのどれだけシフトされるかに依存し、上で説明した双三次フィルタの1つの形をとることができ、

【0101】垂直フィルタリングの後の丸め規則は、次式によって定義される。

$$(25)$$

について、 $\text{shiftV} = \{5, 3, 3, 1\}$ である。シフトの量は、第1ステージフィルタ係数値に起因する拡大の補償に必要な量より小さい(例えば、シフトは、近似双三次フィルタについて6ビット未満である)が、後続フィルタリングの中間結果が中間値のダイナミックレンジ内(例えば、16ビットワードの場合に65536個の可能な値)にとどまることを保証するのに十分である。フルシフトと比較して、この短縮されたシフトでは、補間の第1ステージ(810)の後に、中間ピクセル値(820)の精度が保たれる。中間ピクセル値(820)は、 $\gamma$ ビットのダイナミックレンジを有し、 $\gamma$

は、8ビットより大きい。第1ステージで実行されるシフトの量は、使用可能なビット深さおよび補間フィルタの係数に依存するものとして行うことができる。例えば、本明細書に記載の例示の実施形態では、中間値が、16ビットのワード限界に制限される。

【0103】図6の点P<sub>2</sub>と、0から255の範囲(8ビット)の入力値を検討された。近似三次係数 $[-4 \ 53 \ 18 \ -3]$ を8ビット入力値に適用する中間値の範囲は、フィルタ係数からの拡大要因に起因して、-1785から18105までである(約14.3 10ビット、実施については15ビットまでに丸められる)。中間値に近似三次フィルタ係数(追加の拡大を有する)を適用する後続の水平フィルタリングでは、16ビットダイナミックレンジ外の値が作られ、オーバーフローまたはアンダーフローが引き起こされる可能性がある。従って、中間値は、後続水平フィルタリングで16ビットダイナミックレンジ内の値がもたらされること\*

$$(S+64-R) > 7$$

ここで、Sは、水平フィルタリングされた結果であり、Rは、フレームごとに交互に変える丸め制御値である。第1ステージの丸め規則と同様に、第2ステージの丸め規則には、ステージを交互に変える丸め制御(833)およびビットシフト(834)を用いて丸めが含まれる。第1ステージの延期されたシフトのゆえに、第2ステージでのシフトの量は、通常は、選択された水平フィルタについて通常期待されるものより大きく、所望のダイナミックレンジを有する値を出力するように計算される。

【0105】双三次フィルタリングのすべてのケースで、潜在的に、値が負の補間されたピクセル、または値が範囲の最大値(例えば、8ビット出力では255)より大きい補間されたピクセルが作られる可能性がある。このような8ビット出力値の場合には、エンコーダおよびデコーダが、許容される範囲に収まるように出力値(836)をクリッピングする。具体的に言うと、アンダーフローが、0にセットされ、オーバーフローが、255にセットされる。クランプの後に、補間された8ビット値(838)が出力される。

【0106】図8では、第2ステージのシフトが、7ビット

$$V_1 = (-1x_1 + 9x_2 + 9x_3 - 1x_4) \quad (27)$$

従って、

$$V_1 = (-1a + 9e + 9i - 1m) \quad (28)$$

$$V_2 = (-1b + 9f + 9j - 1n) \quad (29)$$

$$V_3 = (-1c + 9g + 9k - 1o) \quad (30)$$

$$V_4 = (-1d + 9h + 9l - 1p) \quad (31)$$

である。

【0108】R<sub>v</sub>の適当な値を加算した後に、結果を1

ビットだけ右シフトする。第2ステージでは、中間結果V<sub>1</sub>からV<sub>4</sub>が、1/2ピクセルフィルタによって使用★

$$P_2 = (-1V_1 + 9V_2 + 9V_3 - 1V_4) \quad (32)$$

本を保証するのに十分にシフトされる。P<sub>2</sub>について、最初のシフト量は、5ビットであり、シフトされた中間値のダイナミックレンジは、-55から555までである(約9.3ビット、実施については10ビットまでに丸められる)。シフトされた中間値に近似三次フィルタ係数を適用することからの出力の範囲は、-7880から40500までになり、これは、16ビット未満のダイナミックレンジを有する。従って、短縮されたシフトは、16ビットワード限界が完全に利用されるが、補間の第2ステージ(830)中にそれを越えないことが保証されるように計算される。

【0104】第2ステージ(830)は、適当な水平フィルタ(F<sub>H</sub>)を適用して、垂直フィルタによって決定された値(820)からの2次元サブピクセル位置の値を補間する(832)。水平フィルタリングの後の丸め規則は、次の通りである。

$$(26)$$

※ットである。従って、9ビットを有するフィルタリングされた出力値が、保たれる。例えば、P<sub>2</sub>の前の例を続けると、フィルタリングされた出力値の範囲は、-61から316までであり、これは、約8.6ビットのダイナミックレンジを有する(実施については9ビットまでに丸められる)。補間されたデータの有効範囲は、8ビットだけだが、ヘッドルームの余分の1ビットによって、オーバーフロー情報およびアンダーフロー情報が提供される。言い換えると、最上位ビット(すなわち「符号」ビット)がセットされている場合に、アンダーフローまたはオーバーフローがある。具体的に2つのどちらが発生したかは、残りの8つの「仮数」ビットを調べることによって導出される。

【0107】図9から11に、上述し、図8に示した2次元補間をさらに示す。図9に、図6の基準フレーム(600)のサブピクセル位置P<sub>2</sub>(水平に1/2ピクセル、垂直に1/2ピクセル)を示す。2つの1/2ピクセル双三次補間フィルタを使用して、P<sub>2</sub>の値を補間する。第1ステージでは、中間値V<sub>1</sub>からV<sub>4</sub>を、下記の一般形を有する1/2ピクセル双一次フィルタを使用して、近接する整数ピクセル位置から計算する。

★され、P<sub>2</sub>のピクセル値が計算される。具体的に言うと、下記の形を有する1/2ピクセルフィルタが使用される。



【0109】上述したように、第2ステージの結果は、9ビット値を得るために7ビットだけ右シフトされる。この9ビット値には、8つの仮数ビットおよび1つの符号ビットが含まれる。必要なクランプをすべて実行してオーバーフローまたはアンダーフローを補償した後に、最終的な8ビットの補間された値が出力される。

【0110】図10に、図6の基準フレーム(600)\*

$$V_{1, \dots, r} = (-4x_1 + 53x_2 + 18x_3 - 3x_4) \quad (33)$$

このフィルタは、上でP<sub>r</sub>の計算に関して説明したものと同一の形で、基準フレーム(600)の整数ピクセル値に適用される。R<sub>v</sub>の適当な値を加算した後に、結果を3ビットだけ右シフトする。第2ステージでは、中間※

$$P_4 = (-1V_1 + 9V_2 + 9V_3 - 1V_4) \quad (34)$$

第2ステージの結果は、9ビット値を得るために7ビットだけ右シフトされ、必要なクランプがすべて実行され、最終的な8ビットの補間された値が出力される。

【0111】図10には、サブピクセル位置P<sub>4</sub>。(水平に1/4ピクセル、垂直に1/2ピクセル)も示されている。P<sub>4</sub>の値を補間するために、P<sub>4</sub>の補間の技法が、わずかな修正だけを用いて使用される。修正された★20

$$P_4 = (-4V_1 + 53V_2 + 18V_3 - 3V_4) \quad (35)$$

第1ステージおよび第2ステージでのシフトの量は、P<sub>4</sub>を計算する技法と同一である(すなわち、第1ステージのシフトが3、第2ステージのシフトが7である)。

【0112】図11に、図6の基準フレーム(600)のサブピクセル位置P<sub>5</sub>。(水平に1/4ピクセル、垂直に

$$V_{1, \dots, r} = (-4x_1 + 53x_2 + 18x_3 - 3x_4) \quad (36)$$

このフィルタは、P<sub>4</sub>を計算することに関して上で説明したものと同一の形で基準フレーム(600)の整数ピクセル値に適用される。R<sub>v</sub>の適当な値を加算した後、その結果を5ビットだけ右シフトする。第2ステー◆

$$P_5 = (-4V_1 + 53V_2 + 18V_3 - 3V_4) \quad (37)$$

第2ステージの結果が、9ビット値を得るために7ビットだけ右シフトされ、必要なクランプがすべて実行され、最終的な8ビットの補間された値が出力される。

【0113】図9から11には示されていないが、1または両方の次元で3/4ピクセルシフトを有するサブピクセル位置の値も、計算することができる。そのようなサブピクセル位置を計算するために、上で概要を示した方法を、1/4ピクセル双三次フィルタの代わりに適当な3/4ピクセル双三次フィルタを使用することによって修正することができる。

【0114】他の実施形態では、双一次フィルタまたは双一次フィルタと双三次フィルタとの組合せを使用して、サブピクセルサンプリング位置の値を補間する。双一次フィルタの使用によって、双三次フィルタよりも係数を展開しなくてよいので、一般に、実行されるシフトの量(第1ステージの後および全体の後)が減る。例えば、双一次フィルタおよび16ビット中間値を使用する一実施形態では、第1ステージでシフトが実行されず、これ 50

★のサブピクセル位置P<sub>4</sub>。(水平に1/2ピクセル、垂直に1/4ピクセル)を示す。1/4ピクセルおよび1/2ピクセルの双三次補間フィルタを使用して、P<sub>4</sub>の値を補間する。第1ステージでは、中間値V<sub>1</sub>からV<sub>4</sub>を、下記の一般形を有する1/4ピクセル双三次フィルタを使用して、近接する整数ピクセル位置から計算する。

※結果V<sub>1</sub>からV<sub>4</sub>が、1/2ピクセルフィルタによって使用されて、P<sub>4</sub>のピクセル値が計算される。具体的に言うと、下記の形を有する1/2ピクセルフィルタが使用される。

★技法では、第1ステージで1/2ピクセル双三次フィルタを使用して、中間値を決定する。中間ピクセル値の位置は、図10のV<sub>1</sub>からV<sub>4</sub>に示されている。第2ステージ中に、1/4ピクセル双三次フィルタでこの中間値を使用して、P<sub>4</sub>の値を計算する。具体的に言うと、下記の形を有する1/4ピクセル双三次フィルタが使用される。

☆に1/4ピクセル)を示す。2つの1/4ピクセル双三次補間フィルタを使用して、P<sub>4</sub>の値を補間する。第1ステージでは、中間値V<sub>1</sub>からV<sub>4</sub>が、下記の一般形を有する1/4ピクセル双三次フィルタを使用して、近接する整数ピクセル位置から計算される。

◆では、中間結果V<sub>1</sub>からV<sub>4</sub>が、もう1つの1/4ピクセルフィルタによって使用されて、P<sub>4</sub>のピクセル値が計算される。具体的に言うと、下記の形を有する1/4ピクセルフィルタが使用される。

によって16ビットワード境界の使用が最大になり、4ビットの右シフトが最終ステージの後で実行される。同様に、クランプを、最終ステージまで延期することができる。

【0115】上述した方法の基礎となる原理の1つが、所望の「ワードサイズ」境界内にとどまりながら、フィルタリングのすべてのステージで可能な最高の精度を使用することである。出力値が、Dビットのダイナミックレンジを有し、Lビットが、最終ステージで破棄される場合に、フィルタリングの最終ステージの出力は、D+L+1ビットまでを占めることができ、この1つの余分なビットは、アンダーフローおよびオーバーフローのシグナリングに使用される。逆方向に進んで、フィルタリングの最終ステージが、kビットの拡大をもたらす場合に、最終ステージの入力は、D+L+1以内でなければならない。従って、Wビット表現で最大の精度を保つために、下記の関係が存在する。

31

$$D+L+1=W$$

さらに、最終ステージへの入力は、 $D+L-k=W-k$  ビットでなければならぬ。

【0116】この論理を、フィルタリングの終りから2番目のステージに再帰的に適用することができ、以下同様である。実際に、フラクショナルビットを使用して、非2<sup>n</sup> 範囲および拡大要因を表現することによって、上下の限界を狭めることができる。

【0117】図12から図15は、組み合わせで説明したが、複数ステージの補間が別々に適用可能である、さまざまな技法を示す図である。図12から図15には、各々の複数ステージの補間(1200、1300、1400、1500)を、他の複数ステージの補間技法と共に使用することができるさまざまな形で示されていない。

【0118】また、図12から図15のそれぞれに、2つのステージが示されているが、図12から図15に示された複数ステージの補間(1200、1300、1400、1500)技法に、より多くのステージを含めることができる。より一般的には、複数ステージの補間(1200、1300、1400、1500)技法を、

複数の次元の任意の種類の前置可能なフィルタならびにカスケード構造、トレリス構造、または格子構造で実施される任意のフィルタを用いて実施することができる。

【0119】図12から図15に、複数ステージの補間で使用される一般化された入力値、出力値、およびフィルタを示す。第1ステージの入力値、最終ステージの出力値、および中間値のビット深さの特定の選択は、ターゲットのアーキテクチャまたはアプリケーションの技術的仕様に従って任意に拡張することができる。例えば、入力値を、基準フレームの整数ピクセル位置の8ビットピクセル値とし、出力値を、基準フレームのサブピクセル位置の8ビットピクセル値とし、フィルタを、標準の双三次フィルタおよび近似双三次フィルタ(図6から図8に関して上で説明したもの)とすることができる。代替案では、入力値および/または出力値が、異なるビット深さのダイナミックレンジを有するか、または異なるフィルタが使用される。

【0120】それぞれ図4および図5に関して説明したエンコーダまたはデコーダなどのコンポーネントが、複数ステージの補間(1200、1300、1400、1500)を実行することができる。代替案では、別のエンコーダまたはデコーダ、あるいは別のタイプのコンポーネントが、複数ステージの補間(1200、1300、1400、1500)を実行することができる。

【0121】図12に、中間の補間された値に関する拡張されたダイナミックレンジ(ビット単位)を有する複数ステージの補間(1200)の図を示す。第1ステージ(1210)で、コンポーネントが、1または複数の

xビット範囲の入力値(1211)に第1フィルタF、

32

(38)

を適用し(1212)、1または複数のyビット範囲の中間値(1220)を作る。ここで、yはxより大きい。例えば、yビット中間値は、8ビットより広いダイナミックレンジを有するピクセル値であり、xビット入力値は、8ビットのダイナミックレンジを有する。

【0122】詳細には図示されていないまたは1個以上の中間ステージ(1222)のそれぞれで、コンポーネントが、yビット範囲の中間値(1220)にフィルタを適用する。中間ステージからの出力は、1または複数のzビット範囲の中間値(1229)であり、ここで、zはxより大きい(図12から図15では、最終ステージが第2ステージである場合に、第1ステージから出力される中間値が、最終ステージへの入力中間値である)。

【0123】最終ステージ(1230)で、コンポーネントが、1または複数のzビット範囲の中間値(1229)に最終フィルタFを適用する(1232)。最終的な出力は、xビット範囲の出力値(1234)である。複数ステージの補間(1200、1300、1400、1500)技法のそれぞれについて、必要な場合に、コンポーネントは、追加の出力値について複数ステージの補間(1200、1300、1400、1500)を繰り返す。繰り返される補間は、コンポーネントが、前の補間で計算されたある中間値を再利用することができる。

【0124】図13に、スキップされたクランプを用いる複数ステージ補間の技法(1300)の図を示す。クランプの延期によって、例えば、コンポーネントがもたれるか、またはクランプの検出によって、計算が高速になる。延期されたクランプによって、中間値の精度も保たれる。

【0125】第1ステージ(1310)で、コンポーネントが、1または複数のxビット範囲の入力値(1311)に第1フィルタFを適用する(1312)。第1フィルタFの適用の後に、クランプは実行されない。従って、第1フィルタFから出力される1または複数の中間値(1320)が、xビットを超えるダイナミックレンジを有する場合がある。例えば、入力値が、8ビット値であり、第1フィルタFからの出力が、第1フィルタFの係数によって導入される拡張係数(expansion factor)に起因して、9ビットまたはそれ以上のダイナミックレンジを有する。

【0126】詳細には図示されていない0または1個以上の中間ステージ(1322)のそれぞれで、コンポーネントが、1または複数のクランプされていない中間値(1320)にフィルタを適用する。クランプを、0または1個以上の中間ステージ(1322)でスキップすることもできる。0または1個以上の中間ステージ(1322)から出力された中間値(1329)が、最終ス

33

テージ(1330)に入力され、このステージで、コンポーネントが値(1329)に最終フィルタF<sub>1</sub>を適用する(1322)。最終フィルタF<sub>1</sub>からの最終的な出力が、クランプされる(1334)。xビット範囲の値(1336)が出力される。

[0127]図14に、延期されたビットシフトを用いる複数ステージ補間(1400)の図を示す。第1ステージ(1410)で、コンポーネントが、1または複数のxビット範囲の入力値(1411)に第1フィルタF<sub>1</sub>を適用する(1412)。第1フィルタF<sub>1</sub>の適用と共にまたはその後、短縮されたシフト(1414)を実行する。短縮されたシフト(1414)は、xビット範囲の出力値を確保するのに必要なシフトより少なく(第1フィルタF<sub>1</sub>の係数の拡大要因に鑑みて)、従って、第1フィルタF<sub>1</sub>に通常関連するシフトより少ない。

従って、短縮されたシフト(1414)によって、xビットより広いダイナミックレンジ(yビット)を有する1または複数の中間値が作られる。例えば、入力値が、8ビットのダイナミックレンジを有し、中間値が、8ビットを超えるダイナミックレンジを有する。

[0128]詳細には図示されない0または1個以上の中間ステージ(1422)のそれぞれで、コンポーネントが、1または複数の中間値(1420)にフィルタF<sub>2</sub>を適用する。xビット(xビットより大きい)のダイナミックレンジを有する1または複数の中間値(1429)が、0または1個以上の中間ステージ(1422)から出力され、最終ステージ(1430)で、コンポーネントが、値(1429)に最終フィルタF<sub>2</sub>を適用する(1432)。最終フィルタF<sub>2</sub>からの最終的な出力が、最終フィルタF<sub>2</sub>に通常関連するものより多い重たけシフトされ(1434)。これによって、出力値(1436)のダイナミックレンジが、指定されたビット深さに制限される。例えば、出力値(1436)のダイナミックレンジ(ビット単位)が、xまたはx+1と等しい。一実施形態では、第1ステージおよびすべての中間ステージのシフトが、最終ステージで、できる限り延期される。シフトが延期される値は、中間計算に使用可能なビット深さで、めいめいのフィルタの拡大要因に依存する可能性がある。

[0129]図15に、ステージを交互に変える丸め制御を使用する複数ステージ補間技法(1500)を示す。補間技法(1500)の複数のステージは、丸め制御を適用して丸めを調整する形において交互に変える。これは、あるビデオシーケンス内のフレームからフレームへと丸め誤差が累積されるのを防ぐのに役立つ。例えば、低品質ビデオシーケンスに、1次元(パン)または2次元(ズーム)の漸進的な動きが含まれる場合に、丸め誤差の累積によって、フレームからフレームへの漸進的な退色(color fading)がもたらされる可能性がある。これによって、知覚可能な誤差が引き起こされる可

34

能性がある。ステージを交互に変える丸め制御は、そのような退色を防ぐのに役立つ。

[0130]数値例は、右ビットシフトの前にステージを交互に変える丸め制御が適用される丸めを示すのに役立つ。右ビットシフトは、本質的に、右シフトされる値の除算および切捨をもたらす。シフトの前に丸め値を加算することによって、シフトされる値が、必ず切り下げられる(切捨)のではなく、上または下に(最も近い整数に)丸められるようになる。丸め制御を使用することによって、境界の値について丸めの方法(上または下)が変更される。例えば、複数ステージのそれぞれで、フィルタリングの出力が、右シフトの前に右シフトの「除数」の1/2を加算する(例えば、5ビット右シフトの前に $2^4 = 16$ を加算する。7ビット右シフトの前に $2^6 = 64$ を加算する)ことによって調整されると仮定する。この加算の効果は、0.5またはそれ以上の小数成分を有する値が(ビットシフトに対応する除算の後に)(次に大きい整数に)切り上げられることである。このような値は、そうでなければ右シフトによって(次に小さい整数に)切り捨てられる。加算にかかわらず、0.5未満の小数成分を有する値は(ビットシフトに対応する除算の後に)、まだ右シフトによって(次に小さい整数に)切り捨てられる。丸め制御によって、ある境界の値の丸めの方法が変更される。例えば、複数ステージのそれぞれで、フィルタリングの出力が、右シフトの前に0または1(交互に変える丸め制御値)を加算することによって、さらに調整される(例えば、 $2^{x+1} - 1$ または $2^{x+1} - 1 - 1$ )。丸め制御調整の効果は、0.5の小数成分を有する(ビットシフトに対応する除算の後に)値の丸めの方法が変更されることである。1が除算される場合に、そのような境界の値が、切り下げられる。そうでない場合には、そのような境界の値が、切り上げられる。

[0131]複数のステージのそれぞれで、複数ステージ補間の前に、0と1の間で交互に変える丸め制御値が使用され、従って、異なるステージで、丸め制御値が適用される形が交互に変わる。代替案では、複数ステージ補間技法(1500)で、それ自身がステージからステージへ交互に変わる丸め制御値が使用される。

[0132]図15の第1ステージ(1510)で、コンポーネントが、1または複数のxビット範囲の入力値(1511)に第1フィルタF<sub>1</sub>を適用する(1512)。第1フィルタF<sub>1</sub>の適用と共にまたはその後、丸め(1514)が、第1フィルタF<sub>1</sub>からの出力に対して実行される。丸め(1514)は、ステージで交互に変える丸め制御によって調整される。例えば、第1ステージ(1510)で、ステージで交互に変える丸め制御によって、出力値が境界の値(そうでなければ出力値が下にくめられる)である場合に、出力値が最も近い整数に四捨五入に丸められるようになる。1または複数の

丸められた中間値(1520)が、第1ステージから第2ステージ(1530)へ出力される。

【0133】第2ステージ(1530)では、コンポジットが、1または複数の中間値(1520)に第2フィルタF<sub>2</sub>を適用する(1532)。丸め(1534)が、第2フィルタF<sub>2</sub>からの出力に対して実行される。第2フィルタF<sub>2</sub>の適用と共にまたはその後に、丸め(1534)が、ステージで交互に変える丸め制御を用いて実行され、この丸め制御によって、限界の値について第1ステージと反対の方向の丸めが行われるようになる。例えば、第2ステージ(1530)で、ステージで交互に変える丸め制御によって、出力値が限界の値である場合に、出力値が、最も近い整数に向かって下に丸められるようになる。1または複数の中間値(1536)が、第2ステージから出力され、0または1個以上の追加ステージ(1540)でこれらを使用することができる。0または1個以上の追加ステージ(1540)に、さらに、ステージで交互に変える丸め制御を含めることができる。

【0134】交互に変える丸め制御は、連続するステージでの適用に制限されるのではなく、ステージのさまざまな他の組合せで適用することができる。さらに、第1方向を、複数のパラメータに依存するものとすることができる。例えば、ビデオエンコーダまたはビデオデコーダで、第1方向を、前のフレームで使用された丸め制御または補間されるフレームのタイプ(例えば、1フレーム、Pフレーム、またはBフレーム)に依存するものとする。他の実施形態では、第1方向に、臨時情報(casual information)(例えば、過去に符号化/復号化された情報)から暗黙のうちに導出されるか、擬似乱数ジェネレータを使用して導出されるか、ビットストリームの一部としてシグナリングされる、定数をセットすることができる。ステージで交互に変える丸め制御は、単一方向フィルタ、双三次フィルタ、および近似双三次フィルタを含む。さまざまな補間フィルタのどれかを使用して複数ステージ補間に適用することができる。

【0135】(D、クロミナス動きベクトル)クロミナス(クロマ)動きベクトルは、同一位置の輝度動きベクトルから暗黙のうちに導出されるので、その精度は、限られ、単純化のスコープが提供される。この単純化では、コーディングされるビデオの知覚される質を大きく落とさずに、エンコーダおよびデコーダでのクロミナス値のサブピクセル補間の計算の複雑さを減らすことができる。さらに、エンコーダおよびデコーダで、クロミナス動きベクトルの丸めおよび補間の異なるモードの間で切り替えることができる。例えば、あるモードでは、より高い計算的複雑さと引き換えに、コーディングされるビデオの品質を際立たせる。別のモードでは、品質を多少犠牲にして、計算的単純さを際立たせる。

【0136】一実施形態では、ビデオエンコーダおよびビデオデコーダで、シーケンスレベルの1ビットフィールド「FASTUVMC」を使用して、クロミナス値のサブピクセル補間およびクロミナス動きベクトルの丸めを制御する。従って、ビデオエンコーダおよびビデオデコーダは、2つの異なるクロミナス丸めモードすなわち、高速モードと基本モードの1つで選択的に動作する。

【0137】図16に、複数のクロミナス丸めおよび補間モードの間での選択の技法(1600)を示す。例えば、それぞれ図4および図5に関して説明したものなどのビデオエンコーダまたはビデオデコーダが、この技法を実行する。

【0138】ビデオエンコーダまたはビデオデコーダは、1ビットフラグFASTUVMCによって、高速クロミナス動き補償モード(フラグ=1)または基本クロミナス動き補償モード(フラグ=0)のどちらかを示されるかを判定する(1610)。例えば、このフラグは、シーケンスレベルのフィールドであり、ユーザ設定に対応し、エンコーダがコーディングされるビデオのビットストリームに書き込み、デコーダがビットストリームから読み取る。代替案では、エンコーダおよびデコーダが、より多くのビットを使用して、例えば2つより多い使用可能なモード間で選択する代わりに、固定された長さまたは可変長のコードを使用してクロミナス丸めおよび/または補間モードをシグナリングする。あるいは、ユーザ設定に対応するシーケンスレベルのフィールドではなく、切替情報が、ビットストリーム内の他所でシグナリングされ、かつ/または異なる判断基準に従ってセットされる。

【0139】ビデオエンコーダまたはビデオデコーダは、基本モード(1620)または高速モード(1630)でクロミナス動き補償を実行する。基本モード(1620)および高速モード(1630)の動きベクトルの丸めおよび補間の詳細は、一実施形態について以下に示す。代替案では、これらのモードが、異なる実施形態を有する。例えば、以下で説明する高速モード(1630)の実施形態で使用するラックアップテーブルを、特定のハードウェアアーキテクチャの所望の性能レベルをたらすために異なるマッピングに変更するか、または異なる精度の動きベクトルについて動作するように変更する。基本モード(1620)および高速モード(1630)の代わりにまたはこれに加えて、エンコーダまたはデコーダで、クロミナス動きベクトルの丸めおよび補間の他のモードを使用することができる。

【0140】一実施形態では、高速モードで(例えば、クロミナス丸めフラグ=1の場合)、1/4ピクセルオフセットにある(すなわち、1/4ピクセルオフセットおよび3/4ピクセルオフセット)クロミナス動きベクトルが、最も近いフルピクセル位置に丸められ、1

1/2ピクセルオフセットにあるクロミナス動きベクトルが、丸められないままにされ、双一次フィルタリングが、クロミナス補間を使用される。このモードでは、エンコーダおよびデコーダの速度が高くなる。この最適化の動機は、(a) 整数ピクセル位置、(b) 1/2ピクセル位置、(c) 少なくとも1つの座標(xまたはy)について1/4ピクセル位置、および(d) 両方の座標について1/4ピクセル位置にあるピクセルオフセット補間の複雑さの間の大きな相違である。a: b:

c: dの比率は、おおむね1:4:4:7:6:6である\*10

```
//RndTbl[-3]=-1,RndTbl[-2]=0,RndTbl[-1]=+1,RndTbl[0]=0
//RndTbl[1]=+1,RndTbl[2]=0,RndTbl[3]=+1
cmv_x=cmv_x+RndTbl[cmv_x%4];
cmv_y=cmv_y+RndTbl[cmv_y%4];
```

\*8. この高速モードを適用することによって、(a) および(b)を優先することができ、従って、デコーディング時間を削減することができる。これは、クロミナス補間だけについて実行されるので、コーディングおよび品質(特に可視の品質)の消失は、どちらも無視してよい。

【0141】この高速モードでは、丸めの最終的なレベルが、下記のようにクロミナス動きベクトルに対して行われる。

(39)

ここで、cmv\_xおよびcmv\_yは、1/4ピクセル単位のクロミナス動きベクトルのx座標およびy座標であり、%は、剰余(または余り)演算を表し、従って、(x%) = -(x%)と定義される(負の数の剰余は、対応する正の数の剰余の負数と等しい)。従って、cmv\_x(またはcmv\_y)が、4によって割られるときに、クロミナス動きベクトルが、整数オフセットを有する。cmv\_x%4=±2のときに、クロミナス動きベクトルが、1/2ピクセルオフセットを有する。cmv\_x%4=±1または±3のときに、クロミナス動きベクトルが、1/4ピクセルオフセットを有する。再マッピング動作からわかるように、1/4ピクセル位置は、クロミナス動きベクトルを最も近い整数位置に丸めることによって禁止される(1/2ピクセル位置は変更されないままになる)。従って、このモードによって、クロミナス座標が整数ピクセル位置および1/2ピクセル位置に再マッピングされる。双一次フィルタリングを、さらなる高速化のためにこのモードのクロミナス補間のすべてに使用することができる。この高速モード実施形態を、複数の丸めモードの間の選択と組み合わせで説明したが、その代わりに、高速モード実施形態を独立に(すなわち、唯一の可能なモードとして)使用することができる。

【0142】図17は、このクロミナス丸めの第1モードを示す表図(1700)である。第1行(1710)に、1/4ピクセル精度の輝度動きベクトル値が示されている。輝度動きベクトル値は、整数ピクセル位置から的小数オフセットに関して示されているが、これらを、各整数が1/4ピクセル増分を表す整数値(すなわち、0、1/4、1/2、3/4、1ではなく0、1、2、3、4)として表現することができる。第2行(1720)に、クロミナス動きベクトル値が、高速モードでどのように丸められる、その結果、整数ピクセル精度および1/2ピクセル精度を有するようになるかが示されている。

【0143】この実施形態の第2の基本モード(例え

ば、クロミナス丸めフラグ=0)では、1/4ピクセルオフセットにあり、丸め誤差を有しないクロミナス動きベクトルは、1/4ピクセルオフセットのままである。他のサブピクセルオフセットにあるクロミナス動きベクトルは、最も近いフルピクセル位置または1/2ピクセル位置に丸められる。このモードでは、デコーダの速度が、他のモードより遅くなる可能性があるが、クロミナスピクセル値が計算される精度が、高くなる。従って、この基本モードでは、クロミナス座標が、整数ピクセル位置、1/2ピクセル位置、および1/4ピクセル位置に再マッピングされる。上述した双三次フィルタリングまたは双一次フィルタリングを、クロミナス補間を使用することができる。

【0144】図18は、このクロミナス丸めの基本モードを示す図である。第1行(1810)に、1/4ピクセル精度の輝度動きベクトル値が示されている。第2行(1820)に、上で説明した基本モードで対応するクロミナス動きベクトル値がどのように丸められる、その結果、整数ピクセル精度、1/2ピクセル精度、および1/4ピクセル精度を有するようになるかが示されている。他の実施形態では、クロミナス空間が、クロミナス空間の他の分解能に丸められる。

【0145】さまざまな実施形態に関して本発明の原理を説明し、図示したが、これらのさまざまな実施形態を、そのような原理から逸脱せずに配置および詳細において修正できることを理解されたい。例えば、上述した原理および技法は、ビデオエンコーダおよび/またはビデオデコーダでの使用に制限されない。その代わりに、上述した原理および技法は、値が部分的に1または複数の中間値に基づいて計算されるが、分離可能なフィルタが複数の次元で使用される、すべてのコンピューティングの文脈で適用することができる。

【0146】本明細書で説明したプログラム、処理、または方法が、他の形で示されない限り、特定のタイプのコンピューティング環境に関連せず、それに制限されないことを理解されたい。さまざまな種類の汎用または特

定用途のコンピューティング環境を、本明細書に記載の教示による動作と共に使用するか、そのような環境によって本明細書に記載の教示による動作を実行することができる。ソフトウェアで示された実施形態の要素を、ハードウェアで実施することができ、逆も同様である。  
 【0147】本発明の原理を適用することができ多数の可能な実施形態に鑑みて、本発明として、請求項およびその同等物の範囲および趣旨に含まれるすべてのそのような実施形態を請求する。

【図面の簡単な説明】

【図1】従来技術にかかるビデオエンコーダでの動き予測を示す図である。

【図2】従来技術にかかるサブピクセル動き予測および補償での補間に関するサブピクセル位置を示す図である。

【図3】本明細書に記載複数のの実施形態をその中で実施することができる、適切なコンピューティング環境を示すブロック図である。

【図4】本明細書に記載の複数のの実施形態で使用される一般化されたビデオエンコーダシステムを示すブロック図である。

【図5】本明細書に記載の複数のの実施形態で使用される一般化されたビデオデコーダシステムを示すブロック図である。

【図6】サブピクセル動き予測およびサブピクセル動き補償中のピクセル値補間に関する位置を示す図である。

【図7】サブピクセル位置に関する補間されたピクセル値の計算に使用されるピクセル値を有する整数ピクセル位置を示す図である。

【図8】サブピクセル位置の値を補間する2ステージ補間技法を示す図である。

【図9】水平1/2垂直1/2サンプル位置と、そのサンプル位置の値を計算するのに使用されるサブピクセル位置での中間値とを示す図である。

【図10】水平1/4垂直1/2サンプル位置、水平1/2垂直1/4サンプル位置、およびそれらのサンプル位置の値を計算するのに使用されるサブピクセル位置で\*

\*の中間値を示す図である。

【図11】水平1/4垂直1/4サンプル位置と、そのサンプル位置の値を計算するのに使用されるサブピクセル位置での中間値を示す図である。

【図12】拡張されたダイナミックレンジ（ビット単位）中間値を用いる複数ステージ補間技法を示す図である。

【図13】スキップされたクランプを用いる複数ステージの補間技法を示す図である。

【図14】延期されたビットシフトを用いる複数ステージの補間技法を示す図である。

【図15】ステージを交互に交わる丸め制御を使用する複数ステージの補間技法を示す図である。

【図16】複数のクロミナンス丸めおよび補間モードの間での選択の技法を示す流れ図である。

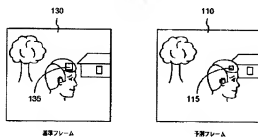
【図17】第1のクロミナンス丸めモードを示す図である。

【図18】第2のクロミナンス丸めモードを示す図である。

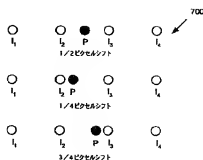
【符号の説明】

- 400 ビデオエンコーダシステム
- 405 現在のフレーム
- 410 動き予測器
- 415 動き情報
- 420 フレームストア
- 425 再構成された前のフレーム
- 430 動き補償器
- 435 動き補償された現在のフレーム
- 445 予測誤差
- 460 周波数変換器
- 466 逆周波数変換器
- 470 量子化器
- 476 逆量子化器
- 480 エントロピー符号化器
- 490 バッファ
- 495 圧縮ビデオ情報

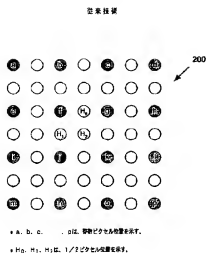
【図1】



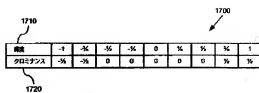
【図7】



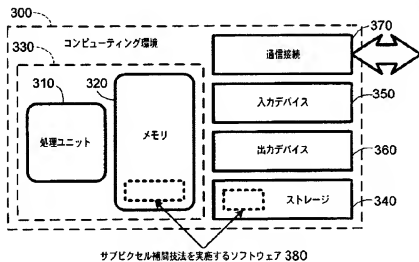
【圖 2】



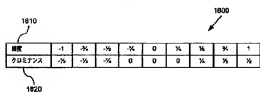
【图 17】



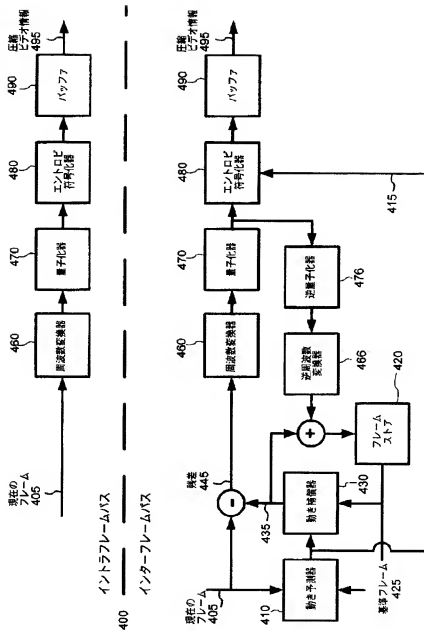
【圖 3】



【圖 18】

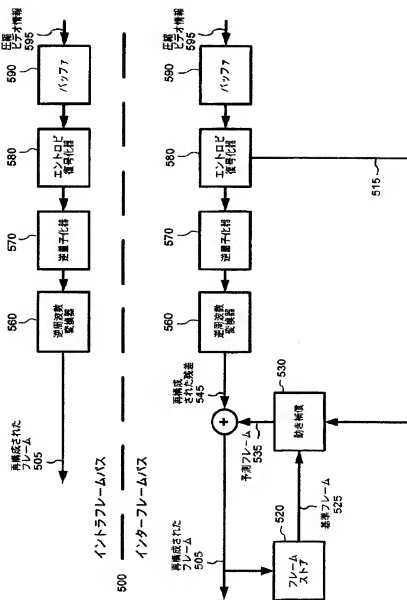


【図4】

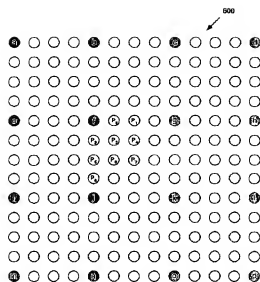




【図5】

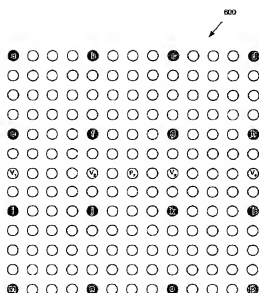


【図6】

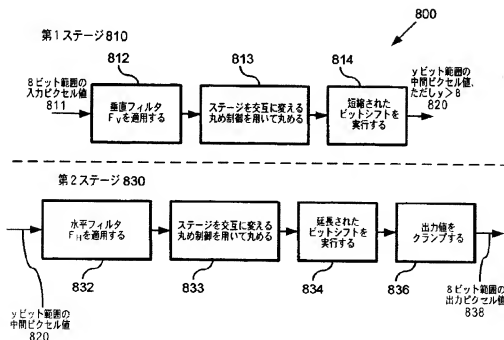


- 円の中や外が、1/4ピクセル範囲のピクセル位置を表す  
 - 0, ..., 15は、参照ピクセル位置を表す  
 -  $P_0, \dots, P_8$ は、異なるサブピクセル位置を表す

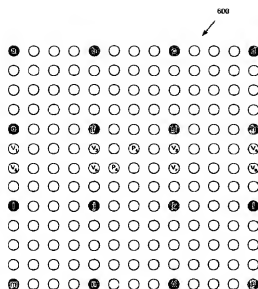
【図9】



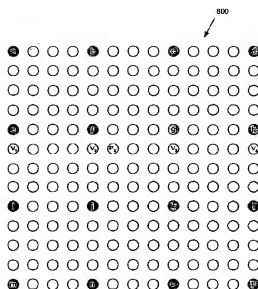
【図8】



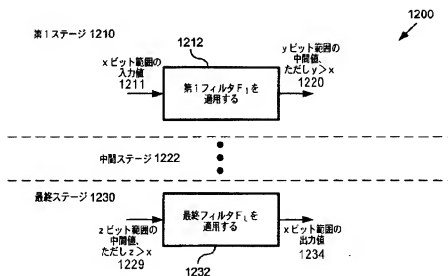
【図10】



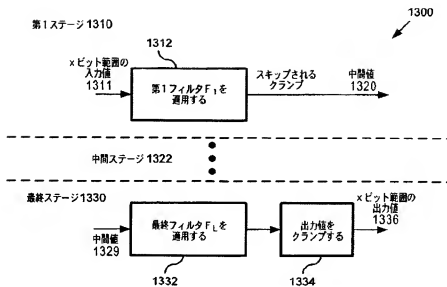
【図11】



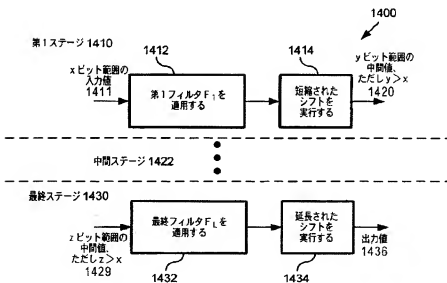
【図12】



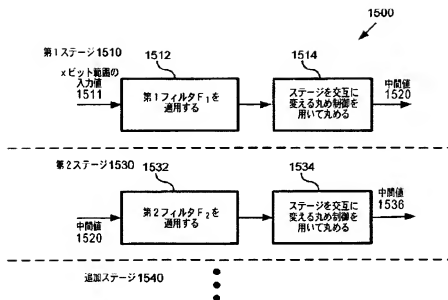
【図13】



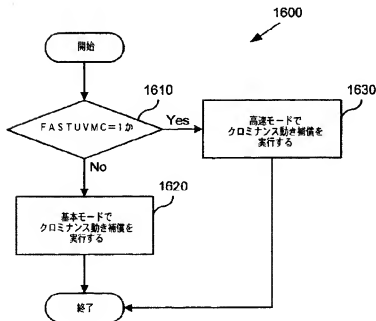
【図14】



【図15】



【図16】



フロントページの続き

(72)発明者 スリドハー スリニバサン

アメリカ合衆国 98109 ワシントン州

シアトル オーロラ アベニュー ノース

1504 ナンバー509

Fターム(参考) 5C059 KC19 LB05 LB18 MA23 MA24

MC11 NN01 NN14 NN21 PP05

PP07 SS20 TA09 TA21 TB04

TC03 TC12 UA15 UA18